METABONOMICS IN THE STUDY OF EYE DISEASES

CHEN LIYAN

B.Sc Chemistry (Hons.), NUS

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHARMACY

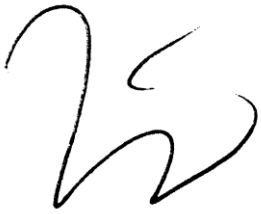
NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



Chen Liyan

21 Jan 2015

Acknowledgements

I wholeheartedly thank Dr Zhou Lei, my supervisor at the Singapore Eye Research Institute, for encouraging me to pursue postgraduate studies. I thank Dr Zhou again, Prof Roger Beuerman and Prof Wong Tien Yin from the Singapore Eye Research Institute for supporting my pursuit of this degree. I thank the Singapore National Eye Centre for sponsorship of my tuition fees through the Health Research Endowment Fund Learning Award.

I thank the postgraduate committee at the department of Pharmacy, chaired by A/P Go Mei Lin, for giving me, a student with an atypical academic background, a chance at postgraduate studies.

I am most grateful to A/P Chan Chun Yong Eric, my supervisor at the University for accepting me into his group and believing in me. I believe I will not find the patience and understanding exemplified by A/P Chan elsewhere, and wouldn't have reached this stage without his guidance.

I thank Prof Ong Choon Nam and A/P Koh Hwee Ling, my thesis committee members for their time and their advice.

I thank Mr Jason Neo and Dr Justin Lim from AB Sciex for their excellent applications support.

I thank my colleague Ms Koh Siew Kwan for assistance with proteomic analysis; Dr Li Jing, Ms Melina Setiawan and Ms Weon Sung Rahn for their assistance with cell

culture; Mr Tian Dechao for assistance with statistical analysis. I thank Dr Hyung Won Choi from the Saw Swee Hock School of Public Health, NUS for discussion and advice on informatics.

I thank Dr Mohammad Kamran Ikram, Dr Cheng Ching-Yu and Prof Wong Tien Yin for access to banked samples from the Singapore Indian Eye Study, and for their input in manuscript preparation.

I thank the senior students and staff from A/P Chan's group, Dr Kishore Kumar Pasikanti, Ms New Lee Sun, Dr Yip Lian Yee and Dr Phua Lee Cheng for the helpful instruction on the GC-MS instrument and advice on data analysis. I also think Dr Chng Hui Ting, Dr Kong Seang Ting, Mr Chan Chun Yip James, Ms Tan Yee Min and all staff and students based at the former S7 corridor for their encouragement and camaraderie.

Table of Contents

Summary

Metabonomics examines the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli. The metabolic phenotype (metabotype) provides unique information that is complementary to genomic, transcriptomic or proteomic profiles. Although metabonomic analysis has been applied in a wide range of medical disciplines such as oncology, neuroscience and toxicology, there are few published studies on eye diseases. This thesis explores the potential of metabonomics in the study of eye diseases and is divided into the following chapters.

1. Characterisation of the Human Tear Metabolome

The tear film is an accessible biological source in studying ocular surface disorders as it is directly associated with the corneal surface. The collection of tears is also a non-invasive procedure that facilitates large-scale clinical studies. A characterisation of the human tear metabolome is essential before the diagnostic potential of the tear film can be exploited. Human tear fluid was analysed with untargeted information-dependent acquisition (IDA) liquid chromatography tandem mass spectrometry (LC-MS/MS). Using isotope pattern-matched peak mining, 60 metabolites were detected in human tears. Among these, 43 metabolites were reported for the first time in tear fluid.

2. Combined Metabonomic and Proteomic Analysis of an *In Vitro* Model of Dry Eye Disease

Most studies that investigated the mechanisms of hyperosmolarity-induced cell damage in dry eye disease using *in vitro* and clinical models were focused on the

analysis of mRNA transcripts and proteins. The metabolome comprises complete set of small molecule metabolites (such as metabolic substrates, intermediates and secondary metabolites) where its fluxes are reflective of downstream level cellular processes. The metabolome could potentially reveal complementary information to the transcriptome and proteome, aiding the understanding of dry eye pathology. Using global metabonomic profiling, glycerophosphocholine was identified as the preferred osmoprotectant in human conjunctiva epithelial cells, over other osmoprotectants such as carnitine and betaine. Increased synthesis of UDP-*N*-acetylglucosamine and activation of the hexosamine pathway were confirmed via proteomic analysis. As *O*-linked protein glycosylation with UDP-*N*-acetylglucosamine is a rapid survival response, modulating glycosylation levels in ocular surface tissues could be important in reducing apoptosis in dry eye disease.

3. Plasma Metabotyping of Diabetic Retinopathy

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and the leading cause of visual impairment in working adults worldwide. Diabetic patients often develop DR despite control of risk factors, suggesting the involvement of the metabolic memory effect. Recent metabonomic studies have uncovered plasma/sera metabolic signatures associated with or predictive of impaired glucose tolerance and diabetes. Given DR is a complex disease where findings from genome-wide association studies have not been conclusive, we postulate that there could be a metabolic signature of DR, and that this signature could be resolvable from diabetes. Plasma samples collected from the Singapore Indian Eye Study were analysed using gas chromatography-time-of-flight mass spectrometry (GC-TOF-MS). 1,5-Gluconolactone, 2-deoxyribonic acid, gluconic acid

and urea were identified as metabolite markers of DR independent of clinical risk factors.

4. Investigation of SWATHTM MS for Metabonomic Analysis

In addition, a short fourth chapter describes the investigation of using a novel mass spectrometry acquisition method, Sequential Window of All Theoretical Fragment Ion Spectra (SWATH^{TM}) MS in metabonomic analysis.

List of Tables

List of Figures

List of Abbreviations

ACN	Acetronitrile
ACR	Albumin-to-Creatinine Ratio
AUC	Area Under Curve
BMI	Body Mass Index
BP	Blood Pressure
CE	Collision Energy
CV	Coefficient of Variation
CXP	Collision Exit Potential
DP	Declustering Potential
DR	Diabetic Retinopathy
eGFR	Estimated Glomerular Filtration Rate
ERK	Extracellular Signal-Regulated Kinase
EI	Electron Ionisation
ESI	Electrospray Ionisation
FASP	Filter-Aided Sample Preparation
FDR	False Discovery Rate
GC-MS	Gas Chromatography-Mass Spectrometry
HbA1c	Blood Glycosylated Haemoglobin
HDL	High Density Lipoprotein
HILIC	Hydrophilic Interaction Chromatography
HMDB	Human Metabolome Database
IDA	Information-Dependent Acquisition
IOBA-NHC	Immortalised Normal Human Conjunctival Epithelial
iTRAQ	Isobaric Tags for Relative and Absolute Quantitation
JNK	c-Jun *N*-terminal Kinase
KEGG	Kyoto Encyclopedia of Genes and Genomes
KFSM	Keratinocyte Serum-Free Medium
LC-MS	Liquid Chromatography-Mass Spectrometry
LC-MS/MS	Liquid Chromatography-Tandem Mass Spectrometry
LDL	Low Density Lipoprotein
LOQ	Lower Limit of Quantitation
MAPK	Mitogen-Activated Protein Kinase
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NIST	National Institutes of Science and Technology
NMR	Nuclear Magnetic Resonance
OSC-PLS-DA	Orthogonal Signal Correction-Partial Least Squares-Discriminant Analysis
PBS	Phosphate-Buffered Saline
PC	Principal Component
PCA	Principal Component Analysis
PUFA	Polyunsaturated Fatty Acid
QC	Quality Control
Q-TOF	Quadrupole-Time of Flight
ROC	Receiver Operating Characteristic
RP	Reverse-Phase
RT	Retention Time
S/N	Signal-to-Noise Ratio
SWATH™	Sequential Window of All Theoretical Fragment Ion Spectra

TOF	Time-of-Flight
UAP1	UDP-*N*-acetylhexosamine Pyrophosphorylase
XIC	Extracted Ion Chromatogram

Papers and Conference Presentations Arising from Work in this Thesis

Chen L, Zhou L, Chan ECY, Neo J, Beuerman RW. Characterization of The Human Tear Metabolome by LC–MS/MS. J Proteome Res. 2011;10(10):4876-82

Chen L, Koh SK, Lei Z. Global Metabonomic Analysis of Human Conjunctival Epithelial (IOBA-NHC) Cells in Response to Osmotic Stress. Singhealth Duke-NUS Scientific Congress; 3 Aug 2012; Singapore (Oral Paper Finalist)

Chen L, Ikram MK, Choi HW, Tian D, Tai ES, Zhou L. GC-MS Metabonomic Profiling of Diabetic Retinopathy. 5th Asia-Oceania Mass Spectrometry Conference; 17 July 2014; Beijing, China (Poster Presenter)

Chen L, Li J, Guo T, Ghosh S, Koh SK, Tian D, et al. Global Metabonomic and Proteomic Analysis of Human Conjunctival Epithelial Cells (IOBA-NHC) in Response to Hyperosmotic Stress. J Proteome Res. (manuscript accepted 11 Aug 2015, DOI 10.1021/acs.jproteome.5b00443)

1. Introduction

1.1. Overview of Metabonomics and Analytical Platforms

The term “metabonomics” was first introduced by Prof. Jeremy Nicholson (1). It is defined as the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli. The related term “metabolomics” on the other hand, referred to the unbiased identification and quantitation of metabolites in a system, with a focus on maximising the coverage of metabolites within an analytical platform (2). The terms “metabonomics” and “metabolomics” are now used interchangeably refer to the global quantitative analysis of metabolites in relation to health and disease (3).

As metabolites are the end products of regulatory processes, the metabolome is a better representation of the phenotype and offers insights complementary to the genome, transcriptome or proteome (3). Thus metabonomics is especially suited to study conditions without strong genetic associations, or conditions with strong environmental contributions.

The main analytical techniques adopted in global metabonomic analysis are nuclear magnetic resonance (NMR) spectroscopy, gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) (4). Both NMR spectroscopy and GC-MS offer high reproducibility and straightforward identification of metabolites (through extensive spectra libraries integrated into

data processing software). In addition, NMR spectroscopy is non-destructive, requires minimal sample preparation and allows absolute quantitative data to be obtained simultaneously. However, NMR spectroscopy is the least sensitive of the three techniques and may be biased towards the detection of high abundance metabolites. Electron ionisation (EI) in GC-MS is highly robust and reproducible, thus enabling the use of spectral libraries in the identification of metabolites. The deconvolution function, a common feature in data processing software, also facilitates spectral resolution of co-eluting metabolites which is not possible by visual inspection of the chromatogram (5). However, the technique may be prone to confounding factors (6) introduced through the sample derivatisation process or thermal degradation of metabolites at elevated temperatures.

More recently, LC-MS has emerged as a complementary metabonomic technique addressing the limitations of NMR spectroscopy and GC-MS. Although LC-MS is highly sensitive and capable of detecting a wide range of metabolites, the technique is susceptible to retention time (RT) drift and matrix effects related to electrospray ionisation (ESI) (7, 8). In addition, ESI also results in the formation of adducts, dimers and in-source fragment product ions (collectively termed redundant signals) alongside $[M+H]^+$/$[M-H]^-$ parent ions, further confounding data analysis and metabolite identification. ESI is also variable among samples analysed on the same instrument in a large batch (9) and fragmentation patterns in tandem mass spectrometry (MS/MS) differ between instruments that apply different mechanisms of ion activation (10). The identification of unknown metabolites detected on LC-MS is limited by the coverage of mass spectra libraries (10).

1.2. Metabonomic Studies in Ocular Fluids and Tissues

While clinical applications of metabonomics in literature encompass a wide range of medical disciplines such as oncology (11), neuroscience (12) and toxicology (13), there are few reports of metabonomics studies on eye diseases. A summary of metabonomic studies on ocular disorders in literature thus far is shown in Table 1.1. The majority of these studies had used ^{1}H-NMR as the analytical platform. From the current database statistics (14) of the Human Metabolome Database (HMDB) (15), over 4500 metabolites have been detected in blood but less than 100 metabolites are known in the aqueous humour and tear fluid. A summary of studies that have profiled the metabolome of ocular fluids and tissues is shown in Table 1.2.

A review of the literature thus far has identified several areas of investigation for the application of metabonomic analysis in studying eye diseases, namely (i) the development of untargeted liquid chromatography-tandem mass spectrometry (LC-MS/MS) platforms for metabonomic applications, (ii) expansion of the known metabolic space for ocular fluids and (iii) application of metabonomic analysis to ocular diseases.

Table 1.1 Metabonomic studies on ocular disorders in literature.

Condition	Sample Type	Analytical Platform	Reference
UV Irradiation	Rabbit Cornea	^{1}H-NMR	(16)
Uveitis	Human Vitreous Humour	^{1}H-NMR	(17)
Dry Eye	Human Reflex Tears	^{1}H-NMR	(18)
Supplements on Dry Eye	Human Reflex Tears	^{1}H-NMR	(19)
Glaucoma Model	Rat Aqueous Humour	^{1}H-NMR	(20)
Diabetic Retinopathy	Human Vitreous Humour	^{1}H-NMR	(21)
Diabetic Retinopathy	Human Plasma	GC-MS	(22)

Table 1.2 Studies profiling the metabolome of ocular fluids and tissues in literature.

Sample Type	Analytical Platform	Number of Identified Metabolites	Reference
Human Aqueous Humour	^{1}H-NMR	13	(23)
Rat Lens	LC-MS (TOF), ^{1}H-NMR	42	(24)
Human Tears, Aqueous Humour	LC-MS/MS (Targeted)	23	(25)

1.3. Thesis Objectives and Scope of Work

1.3.1. Specific Objectives

The objectives of this thesis are:

1. To characterise the metabolites present in human tear fluid using LC-MS/MS.
2. To identify the perturbed metabolites and pathways in an *in vitro* model of dry eye disease
3. To identify the circulating metabolites associated with diabetic retinopathy.
4. To evaluate the performance of SWATH-MS for metabolite quantitation.

1.3.2. Expanding the Known Metabolic Space of Tear Fluid

The tear film is an accessible biological fluid to study disorders of the ocular surface (26). While the tear proteome (27-31) and lipidome (32, 33) have been extensively studied using global proteomic and lipidomic analyses, current knowledge of the tear metabolome is limited to a small number of compounds (25, 34-46). As comprehensive metabonomic profiling has not been performed on tear fluid, the first chapter in this thesis addresses this gap developing untargeted LC-MS/MS methods to identify metabolites present in human tear fluid. Complementary column chemistries and ionisation polarities were explored to increase the coverage of metabolites identified.

1.3.3. Identifying Metabolic Perturbations in an I*n Vitro* Model of Dry Eye

Dry eye disease is a multifactorial disease of the tear film and ocular surface. It is characterised by symptoms of discomfort, disturbance in vision and instability of the tear film. There are two mechanistic subtypes of dry eye disease – aqueous deficient dry eye and evaporative dry eye, and they are known to coexist (47). In particular for evaporative dry eye, tear hyperosmolarity as a metabolic stress, has a role in developing ocular surface inflammation and cell damage (47).

Several *in vitro* studies have investigated the mechanisms associated with hyperosmolarity-induced cell death of ocular surface cells (48-53) by profiling mRNA transcripts or proteins using targeted approaches. Considerably less information is known on the regulation of metabolites in ocular surface cells when exposed to hyperosmotic stress (54). Galbis-Estrada *et al*. had identified a few dry eye-correlated metabolites in tear fluid, but their findings were limited to highly abundant metabolites that could be detected using ^{1}H-NMR (18, 19). Futhermore, the observation that choline levels were elevated in patients with dry eye (18) conflicts with a later report from the authors that choline levels increased after antioxidant supplementation and that this increase in choline levels is associated with a decrease in dry eye symptoms (19).

Studying an i*n vitro* model of dry eye using a more sensitive technique, such as LC-MS/MS may uncover novel information on less abundant metabolites affected in dry eye disease. In the second chapter of this thesis, lysates of human conjunctival cells cultured hyperosmotic media were analysed by LC-MS/MS. Proteomic

analysis were also performed on the cell lysates to identify correlated with dry eye disease.

1.3.4. Investigating Metabolites Correlated With Diabetic Retinopathy

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and the leading cause of visual impairment in working-age adults worldwide (55, 56). The disease is characterised by retinal haemorrhage, aneurysms, neo-vascularisation of the retina. In an early study using ^{1}H-NMR, Barba *et al.* reported decreased vitreous humour concentrations of galactitol and ascorbic acid as metabolic signatures of DR (21). The ideal control group for a nested case-control study on DR are diabetic patients with no signs of DR. However, the controls in Barba *et al.*'s were selected from patients undergoing vitrectomy for macula hole. This selection of the control group was likely limited to the invasive nature of vitreous humour sampling. The obstacles faced in collecting vitreous humour samples limits study replication and the translational potential of any biomarkers identified from vitreous fluid. Recent metabonomic studies have uncovered plasma and sera metabolic signatures associated with, or predictive of impaired glucose tolerance and diabetes (57-66). The lack of clear genetic associations for DR (67) suggest that that metabotyping at the metabolite level might be more suitable for identifying biomarkers of the disease.

When the work in this part of the thesis was being planned, Li *et al*. (22) published a paper showing that the classification of disease severity in DR patients based on metabonomic analysis of plasma metabolites overlapped with the classifications based on traditional chinese medicine. There are several limitations to Li *et al*.'s

study. Firstly, the DR patients in their study had high levels of serum creatinine than the control group, indicating potentially confounding kidney disease, but this was not controlled for in data analysis. Secondly, the fold-changes of their identified marker metabolites (besides glucose) were small and identified only through orthogonal signal correction-partial least squares-discriminant analysis (OSC-PLS-DA) without univariate tests.

In the third chapter of this thesis, plasma metabotypes of DR were investigated using GC-MS with samples from the Singapore Indian Eye Study (68). The Singapore Indian Eye Study is a well-characterised cohort; clinical data such as body mass index (BMI), total cholesterol, low density lipoprotein (LDL) cholesterol and high density lipoprotein (HDL) cholesterol, glycosylated haemoglobin (HbA1c), and blood pressure (BP) were measured using standardised protocols. DR was graded from retinal photographs using a scale modified from the Airlie House classification system. Unsupervised principal component analysis (PCA) was performed to investigate confounding variable-based clustering. Logistic regression was performed on identified metabolite markers of DR to adjust for clinical variables.

1.3.5. Applying SWATH™ MS for Metabolite Quantitation

SWATH™ (Sequential Window of All Theoretical Fragment Ion Spectra) MS is a new data acquisition method (69) that offers multiple reaction monitoring (MRM) -like quantitation without the need to specify target parent and fragment ions during data acquisition. Thus, SWATH™ MS provides superior coverage over MRM by capturing all detectable ions in a sample. Compounds of interest are quantified by

extracting high resolution-acquired product ions from the respective SWATH™ windows.

While SWATH™ MS is increasingly adopted in quantitative proteomic analysis (70-74), there are limited reports on applications of SWATH™ MS for the analysis of small molecule metabolites. There is only one publication each on utilising SWATH™ for toxicological drug screening (75) and identification of drug metabolites (76) respectively. As for endogenous metabolites, Siegel *et al*. had developed SWATH™ MS for the quantitation of *p*-toluenesulfonylhydrazine derivatised aldehydes and ketones (77). This chemoselective derivatising agent produced a signature fragment ion that distinguished metabolites with aldehyde or ketone functional groups from other structures with identical molecular formula.

This chapter of the thesis investigated SWATH™ MS for the quantitation of underivatised endogenous metabolites, including low molecular weight such as betaine (molar mass: 117.15) and sarcosine (molar mass: 89.09) and metabolites with different functional groups.

2. Characterisation of The Human Tear Metabolome

2.1. Introduction to The Tear Film

The tear film is a thin layer of fluid contained between the eyelids, covering the anterior surface of the eyeball. It is also the extracellular fluid that bathes the epithelial cells forming the ocular surface. The tear film provides lubrication, protection and nutrition to the ocular surface (78) and its stability is critical for the optical quality of the eye (79). There are three main layers of the tear film: aqueous, lipid and mucin (78, 80). The aqueous layer is the largest in volume and contains dissolved electrolytes, proteins and metabolites.

The tear film is an accessible biological source in studying ocular surface disorders as it is directly associated with the corneal surface. The collection of tears is also a non-invasive procedure that facilitates large-scale clinical studies. The global profiling of tear metabolites is of particular interest as metabonomics provides a top-down insight into the dynamic biochemical processes occurring in a biological system (81). A characterisation of the human tear metabolome is essential before the diagnostic potential of the tear film can be exploited.

While global-omics platforms have been applied to study the tear proteome (27-31) and lipidome (32, 33) , analysis of the tear metabolome has only been limited to targeted metabolite classes (25, 34-46). The number of metabolites identified in tears prior to this study was significantly lower compared with the metabolic space derived from other biological fluids, such as serum (82) and cerebrospinal fluid

(83). A list of metabolites identified in tear fluid prior to this study is shown in Table 2.1.

NMR spectroscopy and GC-MS were utilised in the characterisation of less-studied metabolomes of saliva (84), amniotic fluid (85) and cerebrospinal fluid (83), these platforms were not sufficiently sensitive for the analysis of tear samples.

Prior to the introduction of qradrupole-time-of-flight (Q-TOF) instruments with faster duty cycles, the typical workflow in untargeted LC-MS metabonomic analysis took on a 'sniper' approach. Only high resolution and accurate TOF-MS would be acquired in the initial profiling experiments (86). Following peak detection and RT alignment, chemometric analysis would be performed to determine differentiating peaks (87), out of which only a fraction can be assigned to be metabolite markers. Metabolites could be putatively identified by performing targeted fragmentation on re-injected samples and comparing the MS/MS spectra obtained with entries in databases and then confirmed by analysing pure standards and comparing their RTs and fragmentation patterns (88). Such downstream data-processing is time-consuming and constitutes the largest bottleneck in a metabonomics workflow. By contrast, data acquisition in untargeted LC-MS-based proteomic analyses are often performed in a shotgun manner (89), with qualitative data acquired simultaneously via IDA. This is enabled by the non-redundant nature of peptide ionisation and the selection of only multiply-charged species for fragmentation. In addition, the use of nano-LC in proteomics further reduces the demands of sensitivity and duty cycle on the instrument.

Table 2.1 Tear metabolites reported in literature thus far.

Metabolite	Reference	Metabolite	Reference
Alkanes and Alkenes		**Carnitines**	
Squalene	(34)	Acetylcarnitine	(41)
		Carnitine	(41)
Amino Acids		Propionylcarnitine	(41)
Alanine	(90)		
Arginine	(90)	**Catecholamines and derivatives**	
Aspartic acid	(90)		
Asparagine	(90)	Epinephrine	(42)
Citrulline	(90)	Norepinephrine	(42)
Cysteine	(35, 36, 90)		
Glutamic acid	(90)	**Hydroxy acids**	
Glutamine	(90)	Ascorbic acid	(35, 36)
Glycine	(90)	Lactate	(37)
Histidine	(90)		
Isoleucine	(90)	**Keto-acids**	
Leucine	(90)	Pyruvate	(37)
Lysine	(90)		
Methionine	(90)	**Nucleotides**	
Ornithine	(90)	Diadenosine triphosphate	(44)
Phenylalanine	(90)	Diadenosine tetraphosphate	(44)
Proline	(90)		
Serine	(90)	Diadenosine pentaphosphate	(44)
Taurine	(90)		
Threonine	(90)		
Tryptophan	(90)	**Peptides**	
Tyrosine	(35, 36, 90)	Glutathione	(35, 36)
Valine	(90)		
		Purines and derivatives	
Amino Ketones		Uric acid	(35, 36, 45)
Urea	(37, 38)		
		Retinoids	
Azoles		Retinol	(46)
Histamine	(43)		
		Steroids and derivatives	
Carbohydrates		Cholesterol	(34)
Glucose	(37, 39, 40)		

This chapter of the thesis developed information-dependent acquisition (IDA) LC-MS/MS methods to analyse the human tear metabolome and identified metabolites through isotope pattern-matched peak mining.

2.2. Methods in Analysis of Human Tear Fluid

2.2.1. Tear Fluid Sample Collection and Preparation

Tears were collected from six healthy volunteers using Schirmer strips without local anaesthesia. All participants had neither ocular complaints nor a history of contact lens use. Written consent was obtained from each participant, and the study was approved by the SingHealth Institutional Review Board (CIRB #2009/910/A). One tear sample was collected from each of the first five volunteers. The sixth volunteer was an in-house participant and twelve tear samples were collected from this participant over a period of 2 weeks between 9 am to noon, with no more than one tear sample collected per eye per day. Tear samples were stored at −80 °C prior to extraction. Metabolites were extracted from the Schirmer strips by vortex-mixing (1200 rpm) each strip in 400 µL of 9:1 methanol/water for 15 min. Each tear extract from the first five volunteers was centrifuged for 10 min at 15,000 g and dried in a vacuum concentrator. Dried samples from the first five volunteers were each reconstituted in 30 µL of water. Tear extracts from the last volunteer were combined into two pools of six extracts each, and dried in a vacuum concentrator. Dried samples from *pool 1* and *2* were reconstituted in 50 µL of water and acetonitrile (ACN) respectively after drying. Schirmer strips with no tears were subjected to the same extraction protocol to serve as blank controls.

2.2.2. IDA LC-MS/MS

Chromatographic separation was performed on a Prominence UFLC system (Shimadzu, Kyoto, Japan). The autosampler and column heater temperatures were maintained at 10 and 40 °C, respectively. The injection volume was 10 µL for all analyses. The fraction of pooled tear samples injected was equivalent to the volume of a single tear sample.

Samples from the first five participants and *pool 1* were injected onto a T3 C_{18} 2.1 × 100 mm, 3 µm reverse-phase (RP) column (Waters Corporation, Milford, MA, USA.). The mobile phase was A: 0.1% formic acid (FA) in water and B: 0.1% FA in ACN. The linear gradient profile was 2% B from 0 to 2 min, 15% B at 6 min, 50% B at 12 min, 95% B from 16 to 18.5 min, and 2% B from 19 to 25 min. The reconstituted sample from *pool 2* was injected onto a ZIC hydrophilic interaction chromatography (HILIC) 2.1 × 100 mm, 3.5 µm column (Merck SeQuant, Darmstadt, Germany). The mobile phase was A: 2 mM ammonium formate in water (pH 6.6) and B: 9:1 ACN: 2 mM aqueous ammonium formate (pH 6.6). The linear gradient profile was 90% B at 0 min to 80% B at 8 min, 10% B from 15 to 18 min, and 90% B from 19 to 24 min. The flow rate for all chromatographic methods was 0.3 mL/min.

Each sample was analysed in positive and negative ionisation polarities using a TripleTOF 5600 Q-TOF mass spectrometer fitted with a DuoSpray ion source (AB Sciex, Concord, Canada). Column effluent was directed to the ESI source. The source voltage was set to 5.0 kV for positive ionisation and 4.0 kV for negative ionisation polarities. The declustering potential (DP) was 80 V and source

temperature was 550 °C for both polarities. The curtain, nebuliser and heater gas flows were set to 25, 45 and 55 units.

Acquisition of MS/MS spectra was controlled by the IDA function of Analyst TF software (AB Sciex). One time-of-flight (TOF) MS survey scan (150 ms) was followed by twenty MS/MS scans (50 ms each) with a total duty cycle time of 1.2 s. The mass range in both TOF and MS/MS scans was *m/z* 50–1000. The following parameters were applied: (i) dynamic background subtraction, (ii) charge monitoring to exclude multiply-charged ions and isotopes, and (iii) dynamic exclusion of former target ions for 5 s. Rolling collision energy (CE) was selected, whereby CE applied was calculated as a function of *m/z* by Analyst TF. Mass accuracy was maintained by the use of an automated calibrant delivery system (AB Sciex) interfaced to the second inlet of the DuoSpray source. Calibrations were performed at the start of a workday or whenever ionisation polarity was changed. The four combinations of column chemistry and ESI polarities are henceforth termed RP (+), RP (−), HILIC (+) and HILIC (−).

2.2.3. Peak Alignment and Multivariate Analysis

MarkerView (AB Sciex) was used to generate *m/z* and retention time (RT) peak lists from the data files of the first five participants analysed with RP (+) and RP (−) analyses. The following parameters were used for peak detection: (i) noise threshold of 20 counts, (ii) minimum chromatographic peak width of 3 scans, (iii) minimum spectra width of 10 mDa, (iv) background subtraction offset of 5 scans and (v) subtraction multiplication factor of 1.2. The peak lists were aligned with the following parameters: RT window of 0.7 to 13 min, RT tolerance of 0.3 min, mass tolerance of 10 ppm, presence of peaks in at least 4 samples and maximum

number of *m/z*-RT peaks at 50000. PCA of aligned peak data was performed using SIMCA (Umetrics AB, Umeå, Sweden) with Pareto scaling.

2.2.4. Isotope Pattern-Matched Peak Mining

Isotope pattern-matched peak mining was performed on pooled tear samples analysed with RP (+), RP (−), HILIC (+) and HILIC (−) using the Extracted Ion Chromatogram (XIC) manager add-on for PeakView (AB Sciex). A summary of the metabolite identification workflow is given in Figure 2.1.

From a list of metabolites derived from the HMDB (15) (2884 entries, MW from 100 to 800), the accurate masses of monoisotopic $[M+H]^+$ and $[M-H]^-$ ions were calculated. The list of $[M+H]^+$ extraction masses was used to extract peaks from the RP (+) and HILIC (+) data files, while the list of $[M-H]^-$ ions was used to extract peaks from RP (−) and HILIC (−) data files. The search window was set to 0 to 10 min and 1 to 21 min for RP and HILIC analyses respectively. Peak detection parameters were set at (i) intensity count over 200, (ii) signal-to-noise ratio (S/N) greater than three and (iii) isotope pattern matching score of 80% and greater.

Entries with peak area values less than 1000 units were removed from the results list. The remaining entries were used to extract peaks from the data files of the blank controls, with the following modifications: (i) adjustment of the search window for each metabolite to ±1 min of their respective retention time as detected in the pooled tear samples, and (ii) deactivation of isotope pattern matching. Entries with peak area values in the pooled tear samples present at less than five times that of the blank controls were removed.

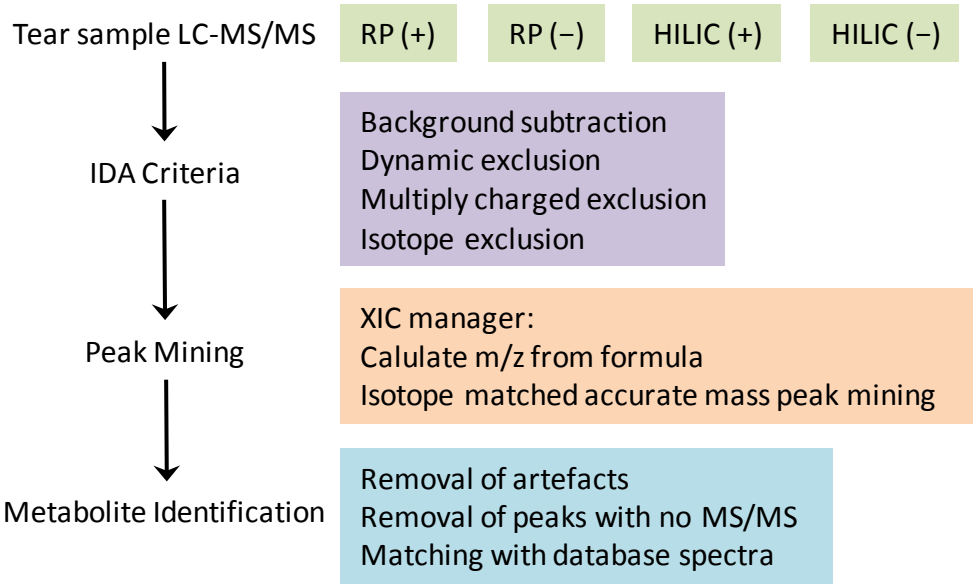


Figure 2.1 Schematic of the IDA-directed LC-MS/MS and isotope pattern-matched peak mining workflow.

Selected ion chromatograms were extracted for each *m/z* in the resultant data table to further screen for isobaric ions at other RTs, as the XIC manager reports only the peak of greatest intensity across the search window. MS/MS spectra of all putative identifications were retrieved from the data files and matched with entries in the METLIN (91), MassBank (92) and HMDB (15) mass spectra databases. Efforts were made to distinguish metabolites from other isobaric compounds whenever possible by virtue of differences in fragmentation pattern. Entries with no experimental MS/MS spectra or corresponding spectra in the databases were not further identified.

Adenosine diphosphate, allantoin, arginine, carnitine, citric acid, glutamine, panthenol, tyrosine, tryptophan, uric acid and xanthine were verified by analysing pure standards with the same LC-MS/MS conditions. Standards were purchased from Sigma Aldrich (St. Louis, MO, USA).

2.3. Results and Discussion: Identification of Tear Metabolites

2.3.1. Resolution of Metabolite Signals from Collection Matrix

The PCA scores plot of tear samples from the first five volunteers and blank controls (collection matrix) derived from RP (+) and RP (−) analyses are shown in Figure 2.2. Both models were optimised at two principal components (PCs) with R^2 values of 0.91. The Q^2 values of the RP (+) and RP (−) models were 0.73 and 0.81 respectively. In both models, tear samples were separated from the blank controls along the first PC. The first PC accounted for 82% of model variation in RP (+) and 87% of model variation in RP (−).

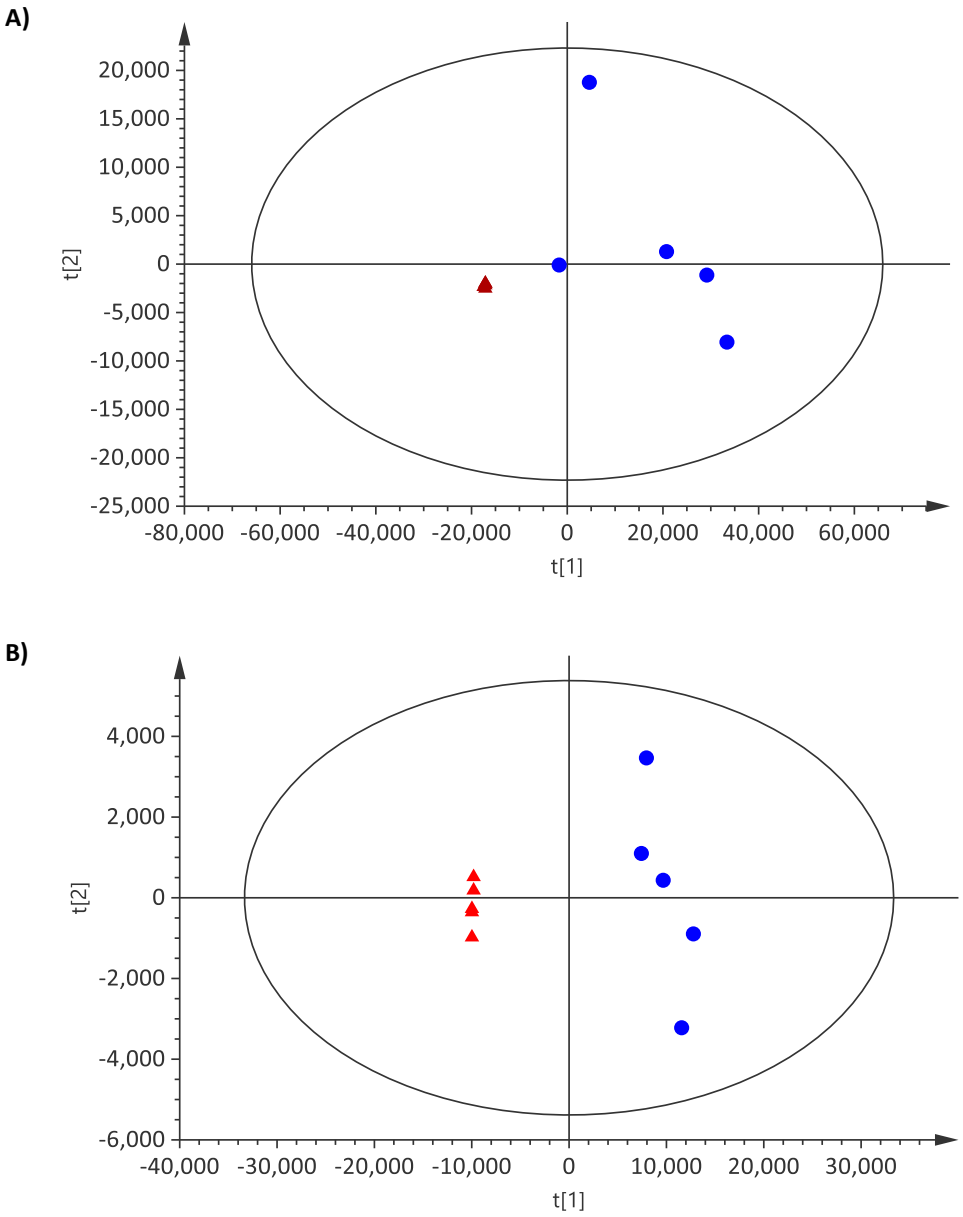


Figure 2.2 PCA scores plot of tear samples (blue circles, n =5) and blank controls (red triangles, n=5) obtained from analysis in A) positive ionisation polarity and B) negative ionisation polarity.

2.3.2. Workings of XIC Manager

Unlike typical data processing in metabonomic analysis where data are subjected to peak detection and experimental masses are obtained, this workflow operates in a reverse manner whereby theoretical masses from the in-house library are used to extract isotope pattern-matched peaks from experimental data. Although a recent publication reported the use of similar software for drug screening in conjunction with targeted analysis (93), there is no precedence of isotope pattern-matched peak mining in global metabonomic analysis.

Representative features of XIC Manager are shown in Figure 2.3, with the identification of xanthine as an example. Figure 2.3A shows the output of XIC manager after isotope pattern-matched peak mining was performed on the RP (+) data file. The green circles indicate that the errors in mass and isotope pattern were less than 10 ppm and 20% respectively (user-set thresholds). In addition to an output table, XIC manager extracts ion chromatograms of all putative metabolites. The XIC of xanthine (RT of 2.05 min on RP) is highlighted in Figure 1.3B. The TOF-MS and MS/MS spectra of xanthine are shown in Figure 1.4A and Figure 1.4B respectively. The regions highlighted in grey indicate a match to the theoretical isotope distribution.

A)

XIC Manager

#	Name	Formula	Adduct	Extraction Mass (Da)	Found At RT (min)
69	Adenosine	C10H13N5(	H+	268.10403	3.28
70	Guanine	C5H5N5O	H+	152.05669	3.78
71	Guanosine	C10H13N5(	H+	284.09895	3.78
72	Thymidine	C10H14N2(	H+	243.09755	12.87
73	Thymine	C5H6N2O2	H+	127.0502	2.6
74	Uracil	C4H4N2O2	H+	113.03455	2.17
75	Uridine	C9H12N2O	H+	245.07681	15.2
76	Xanthine	C5H4N4O2	H+	153.0407	2.05
77	Kynurenine	C10H12N2(	H+	209.09207	10.29
78	Melatonin	C13H16N2(	H+	233.12845	9.79
79	L-NMMA	C7H16N4O	H+	189.1346	12.81

Sample: tears_strip_pos.wiff (sample 1) Control:

B)

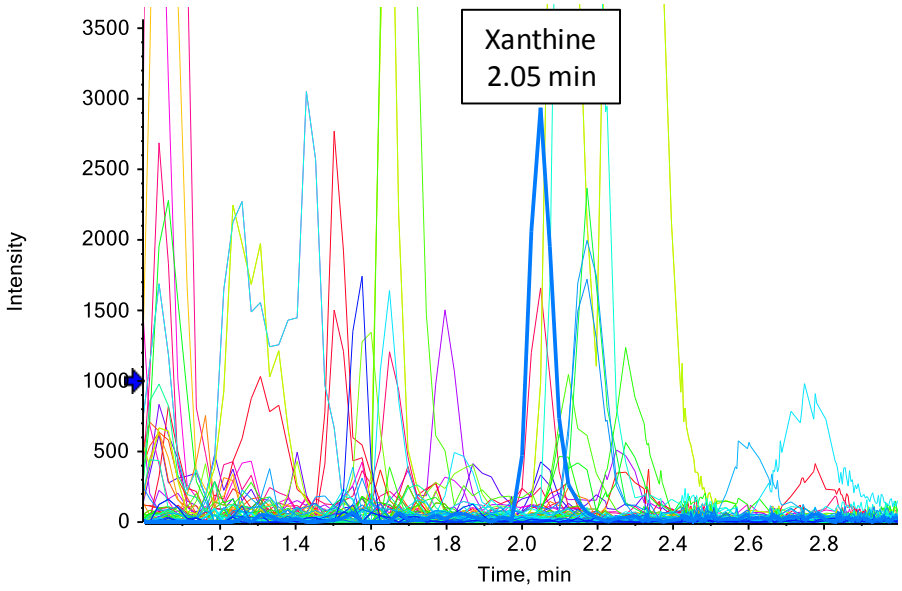
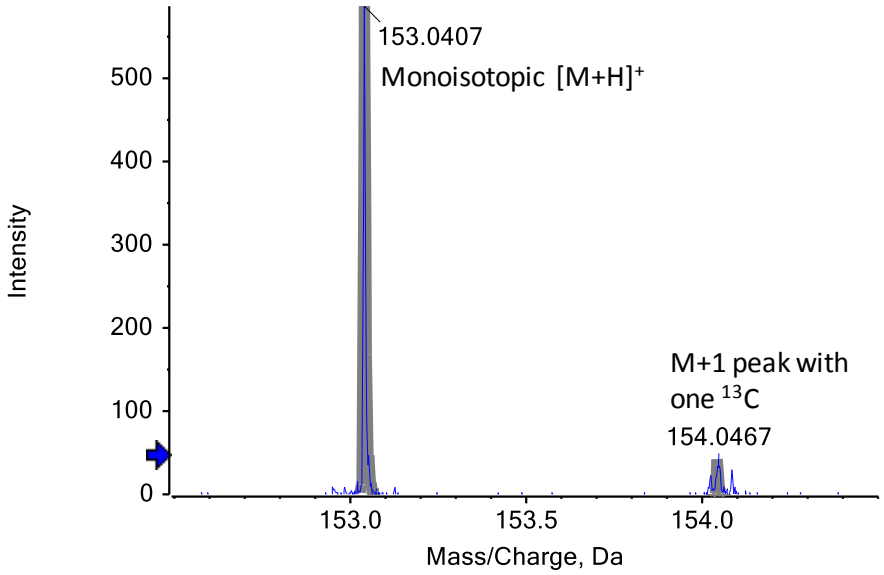


Figure 2.3 A) Output of XIC Manager for RP (+) analysis and B) XIC of xanthine highlighted among other peaks.

A)



B)

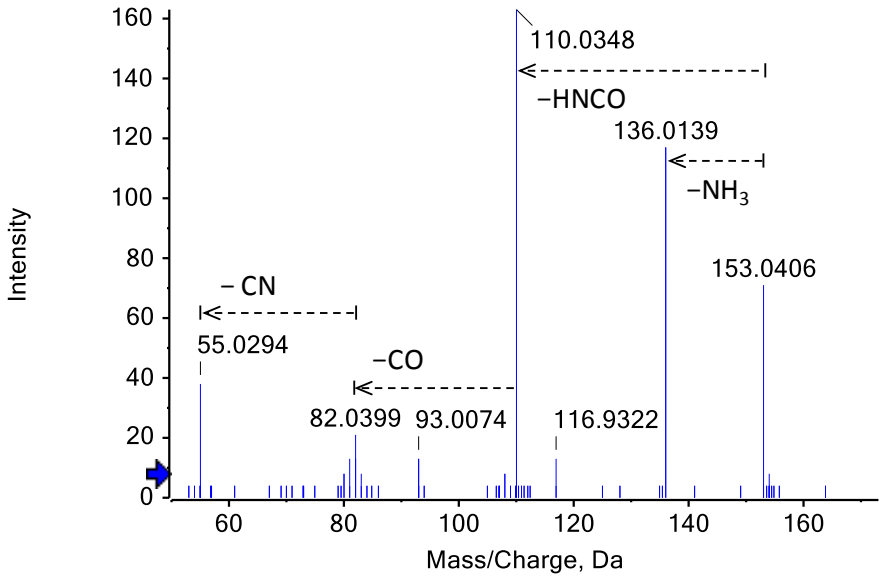


Figure 2.4 A) TOF-MS spectrum of xanthine and B) MS/MS spectrum of Xanthine (CE 10 V).

2.3.3. Identified Tear Metabolites

Using isotope pattern-matched peak mining, 60 metabolites were identified in human tears (Table 2.2). This study reports the presence of 44 of these metabolites in human tear fluid for the first time (Table 2.2) A diverse range of compound classes are represented (amino acids, carnitines, nucleotides, phospholipids etc.), thereby demonstrating good coverage of the analytical method. The RTs of metabolites found in each analytical mode are given in Table 2.3. Metabolites were reported together as a single entry when resolution of isobaric compounds was not possible. For example, fumaric acid and maleic acid, the *trans* and *cis* isomers of butenedioic acid produce identical MS/MS fragments and were not resolved with the chromatography method used in this study. Fumaric acid is a key intermediate in the citric acid cycle and is thus expected to be present in tear fluid. While maleic acid is not an intermediate in major metabolic pathway, it is found in human urine (94) and thus cannot be excluded from being present in tear fluid.

Metabolites that covered a range of chemical structures, polarities and isotope distribution patterns were selected for confirmation. They were adenosine diphosphate, allantoin, arginine, carnitine, citric acid, glutamine, tyrosine, tryptophan, panthenol, uric acid and xanthine.

Table 2.2 Metabolites detected in tears by IDA LC-MS/MS peak mining.

Amino Acids	Carbohydrates	Nucleotides
1-Methylhistidine/ 3-Methylhistidine	*N*-Acetylneuraminic acid	Guanosine monophosphate
Arginine [a, *]	Carnitines	Inosine monophosphate
Asymmetric dimethylarginine/ Symmetric dimethylarginine	Acetylcarnitine [c]	Uridine diphosphate
	Carnitine [c,*]	Uridine monophosphate
	Hexanoylcarnitine	UDP-*N*-acetylgalactosamine/ UDP-*N*-acetylglucosamine
	Palmitoylcarnitine [c]	
Citrulline [a]		
Creatine	Cyclic Amines	
Glutamic acid [a]	Niacinamide	
Glutamine [a, *]		Peptides
Homoarginine	Dicarboxylic Acids	Oxidised glutathione
Hydroxyproline	Fumaric acid/ Maleic acid	
Phenylalanine [a]		Phospholipids
Proline [a]		1-Palmitoyl-lysophosphatidylcholine
Pyroglutamic acid	Nucleosides	
Serine [a]	1-Methyladenosine	
Taurine [a]	Adenosine	Purines and derivatives
Threonine [a]	Cytidine	Hypoxanthine
Tyrosine [b, *]	Guanosine	Theobromine
Tryptophan [a, *]	Inosine	Uric acid [d, *]
Urocanic acid	*S*-Adenosyl-homocysteine	Xanthine*
Valine [a]		
	S-Adenosyl-methionine	Pyridoxals and derivatives
Amino Alcohols	Uridine	
Panthenol*	Xanthosine	4-Pyridoxic acid
Amino Ketones	Nucleotides	Quaternary Amines
Allantoin*	Adenosine diphosphate*	Acetylcholine
Creatinine	Adenosine monophosphate	Glycerolphosphocholine
		Phosphocholine
Aromatic Acids	Cytidine monophosphate	
Cinnamic acid	Cytidine diphosphate choline	Tricarboxylic Acids
o-Coumaric acid/ *m*-Coumaric acid/ *p*-Coumaric acid		Citric acid*

[a] reported in (90)
[b] reported in (35, 36, 90)
[c] reported in (41)
[d] reported in (35, 36, 45)
* verified with standard

Table 2.3 Formulae, monoisotopic masses and RTs of metabolites identified using multiple analytical modes.

#	Formula	Mass (*m/z*)	Name	RT (min)			
				RP (+)	RP (−)	HILIC (+)	HILIC (−)
1	$C_{11}H_{15}N_5O_4$	281.1124	1-Methyladenosine	1.67			
2a	$C_7H_{11}N_3O_2$	169.0851	1-Methylhistidine [a]			10.80	
2b			3-Methylhistidine[a]				
3	$C_{24}H_{50}NO_7P$	495.3325	1-Palmitoyl-lysophosphatidylcholine			1.76	
4	$C_8H_9NO_4$	183.0532	4-Pyridoxic acid		2.67		
5	$C_9H_{17}NO_4$	203.1158	Acetylcarnitine	1.53		5.13	
6	$C_7H_{15}NO_2$	146.1181	Acetylcholine			10.09	
7	$C_{10}H_{13}N_5O_4$	267.0968	Adenosine	3.29	3.13	1.42	1.41
8a	$C_8H_{18}N_4O_2$	202.1430	Asymmetric dimethylarginine[b]			16.26, 16.49	
8b			Symmetric dimethylarginine[b]				
9	$C_{10}H_{15}N_5O_{10}P_2$	427.0294	Adenosine diphosphate*		1.28		
10	$C_4H_6N_4O_3$	158.0440	Allantoin[c,*]				2.43
11	$C_{10}H_{14}N_5O_7P$	347.0631	Adenosine monophosphate	1.55	1.60	3.70	3.68
12	$C_6H_{14}N_4O_2$	174.1117	Arginine*			17.08	
13	$C_7H_{15}NO_3$	161.1052	Carnitine*	1.01		7.21	
14	$C_9H_8O_2$	148.0524	Cinnamic acid				2.01
15a	$C_6H_8O_7$	192.0270	Citric acid*		1.83		
16	$C_6H_{13}N_3O_3$	175.0957	Citrulline			6.79	
17	$C_9H_{14}N_3O_8P$	323.0519	Cytidine monophosphate		1.13		

#	Formula	Mass (*m/z*)	Name	RT (min)			
				RP (+)	RP (−)	HILIC (+)	HILIC (−)
18a	$C_9H_8O_3$	164.0473	*o*-Coumaric acid[a]	2.12		3.22	
18b			*m*-Coumaric acid[a]				
18c			*p*-Coumaric acid[a]				
19	$C_4H_9N_3O_2$	131.0695	Creatine	1.06		5.39	
20	$C_4H_7N_3O$	113.0589	Creatinine				1.78
21	$C_9H_{13}N_3O_5$	244.0935	Cytidine	1.51			
22	$C_{14}H_{26}N_4O_{11}P_2$	488.1073	Cytidine diphosphate choline			10.05	
23a	$C_4H_4O_4$	116.0110	Fumaric acid[a]		1.18		
23b			Maleic acid[a]				
24	$C_5H_9NO_4$	147.0532	Glutamic acid			5.56	5.42
25	$C_5H_{10}N_2O_3$	146.0691	Glutamine*		1.03	5.69	5.66
26	$C_8H_{20}NO_6P$	257.1028	Glycerophosphocholine			8.13	
27	$C_{10}H_{14}N_5O_8P$	363.0580	Guanosine monophosphate	1.65	1.73		
28	$C_{10}H_{13}N_5O_6$	284.0917	Guanosine	3.79	3.76	2.35	2.35
29	$C_{13}H_{25}NO_4$	259.1784	Hexanoylcarnitine			2.43	
30	$C_7H_{16}N_4O_2$	188.1273	Homoarginine			16.75	
31	$C_5H_9NO_3$	131.0582	Hydroxyproline			4.64	4.63
32	$C_5H_4N_4O$	136.0385	Hypoxanthine	1.80	1.80	1.55	
33	$C_{10}H_{13}N_4O_8P$	348.0471	Inosine monophosphate	1.70	1.80		
34	$C_{10}H_{12}N_4O_5$	268.0808	Inosine	3.79	3.76	1.84	1.83
35	$C_{11}H_{19}NO_9$	309.1060	*N*-Acetylneuraminic acid				4.58
36	$C_6H_6N_2O$	122.0480	Niacinamide	1.65			

#	Formula	Mass (*m/z*)	Name	RT (min)			
				RP (+)	RP (−)	HILIC (+)	HILIC (−)
37	$C_{20}H_{32}N_6O_{12}S_2$	612.1520	Oxidised glutathione	2.15	2.29		
38	$C_{23}H_{45}NO_4$	399.3349	Palmitoylcarnitine			1.64	
39	$C_9H_{19}NO_4$	205.1314	Panthenol*	5.66	5.69	1.18	
40	$C_9H_{11}NO_2$	165.0790	Phenylalanine	4.37	4.35	2.01	
41	$C_5H_{14}NO_4P$	183.0661	Phosphocholine	1.04			
42	$C_5H_9NO_2$	115.0633	Proline				3.80
43	$C_5H_7NO_3$	129.0426	Pyroglutamic acid			5.66	
44	$C_{14}H_{20}N_6O_5S$	384.1216	*S*-Adenosyl-homocysteine	2.17			
45	$C_{15}H_{22}N_6O_5S$	398.1372	*S*-Adenosyl-methionine			15.71	
46	$C_3H_7NO_3$	105.0426	Serine			6.01	5.98
47	$C_2H_7NO_3S$	125.0147	Taurine		0.98	4.56	4.56
48	$C_7H_8N_4O_3$	181.0647	Theobromine	5.98			
49	$C_4H_9NO_3$	119.0582	Threonine				4.76
50	$C_{11}H_{12}N_2O_2$	204.0899	Tryptophan*	6.65	6.68	2.28	2.28
51	$C_9H_{11}NO_3$	181.0739	Tyrosine*	2.15	2.22		3.24
52	$C_9H_{14}N_2O_{12}P_2$	404.0022	Uridine diphosphate			7.48	
53	$C_9H_{13}N_2O_9P$	324.0359	Uridine monophosphate		1.33		3.78
54	$C_5H_4N_4O_3$	168.0283	Uric acid*	1.72	1.78	3.12	2.94
55	$C_9H_{12}N_2O_6$	244.0695	Uridine	2.17	2.32		1.54
56a	$C_{17}H_{27}N_3O_{17}P_2$	607.0816	UDP-*N*-acetylgalactosamine[a]			6.26	
56b			UDP-*N*-acetylglucosamine[a]				
57	$C_6H_6N_2O_2$	138.0429	Urocanic acid	1.55	1.58	1.96	1.93

#	Formula	Mass (*m/z*)	Name	RT (min)			
				RP (+)	RP (−)	HILIC (+)	HILIC (−)
58	$C_5H_{11}NO_2$	117.0790	Valine			3.11	
59	$C_5H_4N_4O_2$	152.0334	Xanthine*	2.05	2.17		1.51
60	$C_{10}H_{12}N_4O_6$	284.0757	Xanthosine		4.52		

[a] chromatographically and spectrally unresolvable
[b] partially resolved by chromatography, spectrally unresolvable
[c] identified through manual search
* verified with standard

2.3.4. Evaluation of Workflow

In this study, entries with peak intensities in tear samples that were less than five times the levels in blank controls were removed. This filter significantly reduced the likelihood of mis-identifying artefacts leeched from Schirmer strips as endogenous tear metabolites. These results show that the LC-MS/MS platform is sufficiently sensitive in detecting endogenous metabolites from the small volume of tears in the presence of matrix components.

Coverage of the tear metabolome was increased by the utilisation of complementary column chemistries and ionisation in both positive and negative polarities. The sole use of RP (+), RP (−), HILIC (+), or HILIC (−) would result in the identification of only 26, 25, 35 or 22 metabolites respectively (Table 2.3). While HILIC is accepted to be complementary to RP in selectivity, (95) a total reversal in the elution order of metabolites should not be expected. For example, hypoxanthine eluted before inosine on both RP and HILIC columns in our study and this elution pattern has also been reported in literature (96, 97). The HILIC particle is zwitterionic, thus its selectivity and retention are influenced by pH and ionic strength to a greater extent than RP.

The acquisition of both TOF-MS and MS/MS spectra in a single data file was made possible by utilising a short duty cycle with a high number of product ion scans (20 MS/MS scans in a cycle time of 1.2 s). This facilitated the mass resolution of co-eluting metabolites and increased the chances of $[M+H]^+$/$[M-H]^-$ species being selected for fragmentation amongst redundant signals. In addition, exclusion of former selected target ions in IDA parameters improved the instrument's

productivity in acquiring non-repeated MS/MS spectra by selecting new-eluting metabolite signals for fragmentation.

Mass accuracy and high mass resolution form the foundation of this peak mining workflow. These parameters were measured to be 6 ppm (average 2.4 ppm from all identified metabolites) and over 30,000 at *m/z* 400 respectively (Figure 2.5). The use of an orthogonal isotope pattern filter greatly reduced the probability of detecting background noise and redundant signals as monoisotopic peaks, and eliminated false positives from compounds with similar molecular weights (98).

Several challenges were encountered during metabolite identification: i) metabolites with identical molecular formulae, ii) in-source fragmentation and iii) adduct formation and these are illustrated henceforth. In the first example, a peak with *m/z* 282.083 was detected at 3.76 min on RP (+), with guanosine or 8-hydroxy-deoxyguanosine being possible identifications. By comparing the experimentally-derived MS/MS spectra of this peak to MS/MS spectra of guanosine and 8-hydroxydeoxyguanosine entries in METLIN (91), the peak was identified to be guanosine.

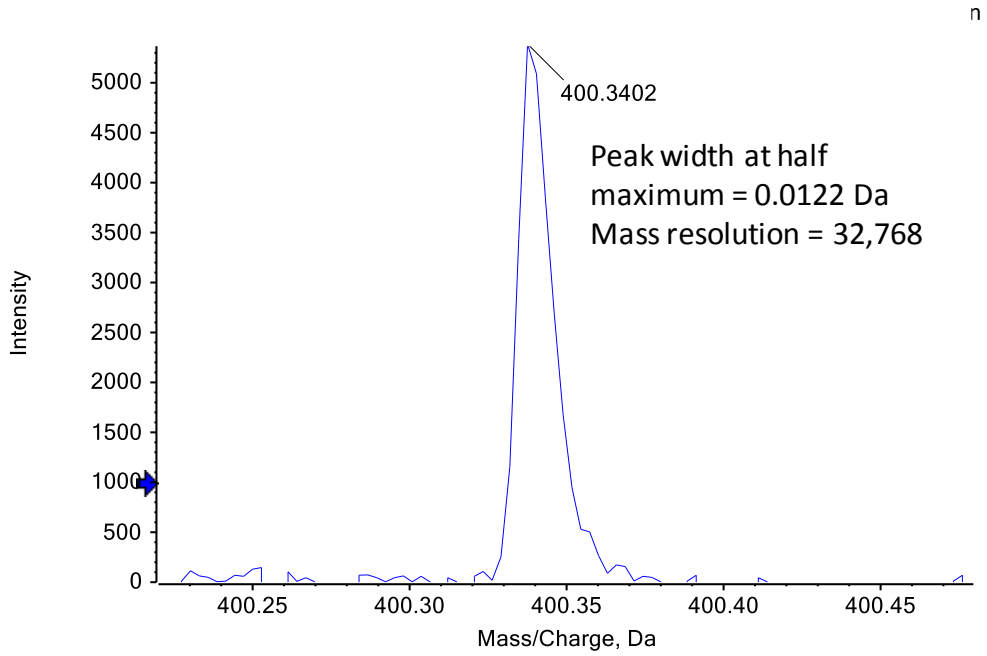


Figure 2.5 TOF-MS spectrum of palmitoylcarnitine.

In the case of in-source fragmentation, inosine (*m/z* 269.088) and hypoxanthine (*m/z* 137.046) were both putatively identified in XIC manager at the same retention time of 3.79 min on RP (+), with their MS/MS spectra matching the respective entries in METLIN (91) and MassBank (92). Inosine has been reported to fragment at source to form hypoxanthine (99) should elute at a later RT than hypoxanthine. Extraction of *m/z* 137.046 across the whole chromatogram showed another peak at 1.80 min. This peak at 1.80 min was identified to be hypoxanthine based on its MS/MS spectra. In the last example, *N*-succinyl-2-amino-6-ketopimelate (*m/z* 290.087) was putatively identified at 3.29 min but a query of the compound on HMDB (15) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (100) showed that *N*-succinyl-2-amino-6-ketopimelate is not involved in primary metabolism in humans. As this ion of *m/z* 290.087 showed identical retention with adenosine and a difference of 22.990 Da from the monoisotopic mass, this peak was identified as the sodium adduct of adenosine.

The efficiency of isotope pattern-matched peak mining was evaluated using the subset metabolite candidates with molecular weight between 100 to 300 data searched on the RP (+) data file (Table 2.4). After removing entries with no MS/MS spectra and duplicate *m/z* values, 36 unique extraction masses remained. Out of these, 20 metabolites were identified, and 3 entries failed to match with database spectra. Five entries were found to be adducts or product ions from in-source fragmentation (n = 5) of other metabolites. Eight entries could not be identified as there were no mass spectra of these candidates in databases. Thus, the false discovery rate the isotope pattern-matched peak mining workflow prior to MS/MS identification is less than 10% (3 out of 36).

The accuracy of isotope abundances is influenced by several factors such as ion intensity, relative abundance of isotope peak and competitive ionisation (101). This workflow was based on the presumption that at least the M+1 isotope of each metabolite was detectable on the Q-TOF instrument, and the intensity of the respective isotopes were within 80% accuracy of calculated isotope distributions.

Given these limitations, the list of identified metabolites presented in Table 2.2 is by no means exhaustive as metabolites displaying the following attributes would possibly remain unreported: (i) strong adduct formation with sodium or formate ions, (ii) present at a low concentration in tear samples, and (iii) low molecular weight with a small number of carbon atoms. For example, allantoin ($C_4H_6N_4O_3$) was only identified upon further investigation. It was initially unreported by XIC manager as the M+1 isotope peak (calculated relative abundance of 5% to the monoisotopic peak) was below the detection limit of the mass spectrometer.

Table 2.4 Evaluation of workflow using subset of metabolite candidates with molecular weight from 100 to 300 searched on RP (+)

Metabolite assignment	Number of metabolite candidates
Starting entries	109
Entries with no MS/MS spectra	13
Duplicate m/z	59
Putative identifications	36
Identified metabolites	20
No match to database entries	3
Adducts and fragments	5
No spectra in databases	8

2.4. Summary for Characterisation of The Human Tear Metabolome

Isotope pattern-matched peak mining enabled the identification of 60 metabolites in human tear fluid. These metabolites encompass a wide range of compound classes, and 44 were identified for the first time in tear samples. The separation of tear samples from Schirmer strip blanks by PCA shows that the LC-MS/MS analytical platform is sufficiently sensitive for global metabonomic profiling of tear fluid to study ocular diseases.

3. Metabonomic and Proteomic Analysis of an *In Vitro* Model of Dry Eye Disease

3.1. Introduction to Dry Eye Disease

Dry eye disease is a multifactorial disease of the tear film and ocular surface. It is characterised by symptoms of discomfort, disturbance in vision and instability of the tear film. There are two mechanistic subtypes of dry eye disease – aqueous deficient dry eye and evaporative dry eye, and they are known to coexist (47). In particular for evaporative dry eye, tear hyperosmolarity as a metabolic stress, has a role in developing ocular surface inflammation and cell damage (47). The prevalence of dry eye is high and estimated to be between 11 to 17% among persons over 40 years old in predominantly-Caucasian populations (102-104) and between 21 to 33% in Asian populations (105-107). The prevalence of dry eye disease increases with age and is higher in females (102).

The disparity in global prevalence rates could be attributed to the use of different diagnostic criteria, with individual clinical signs demonstrating poor agreement with disease severity (108). A combination of clinical assessments such as self-reported symptoms, corneal and conjunctiva staining, tear break up time, tear volume by Schirmer's test and meibomian gland function are used for the diagnosis of dry eye disease and response to treatment (47). While the combination of such methods yields improved sensitivity and specificity (109), diagnosis remains time- and resource-consuming, and highly dependent on the judgement of trained physicians (108). The lack of objective assessment of

treatment outcomes hinders agreement among clinicians and also the development of targeted pharmacotherapies. Artificial tear formulations are available over-the-counter but they merely provide symptomatic relief. In view of the gaps in clinical knowledge, current research efforts in dry eye disease are directed towards (i) selective biomarker discovery and (ii) elucidation of disease mechanism in animal and cell models. Several studies have been conducted to identify biomarkers of dry eye disease using proteomic approaches (27, 110, 111), including subsets of mild dry eye (110) and contact lens-related dry eye (111).

Hyperosmolarity is known to exert pleiotropic effects on cellular function on mammalian renal cells (112), including cell cycle arrest and increased apoptosis (113). Several *in vitro* studies have been performed to investigate the mechanisms associated with hyperosmolarity-induced cell death of ocular surface cells. Li *et al*. were the first to demonstrate the involvement of mitogen-activated protein kinase (MAPK) signalling pathway (48). Subsequent studies have shown that c-Jun *N*-terminal kinases (JNK) and extracellular signal-regulated kinase (ERK) pathways (49) induce apoptosis in human corneal epithelial through a cytochrome c-mediated pathway (50) dependent on transglutaminase (52). Collectively, these *in vitro* mechanistic studies confirmed the involvement of cellular signalling pathways in hyperosmolarity-induced cell death and suggested the potential dysregulation of signalling pathways and endogenous metabolites in dry eye disease.

Most published work have focused on profiling mRNA transcripts or proteins using targeted approaches (48-53). Considerably less information is available on the regulation of metabolites in ocular surface cells when exposed to osmotic stress (54). We postulate that global non-targeted profiling of the metabolome and

proteome could reveal novel and complementary information on pathways affected in dry eye disease. This chapter of the thesis investigated the perturbations in a cell model of dry eye using global metabonomic and proteomic analyses.

3.2. Methods in Analysis of Cell Lysates

3.2.1. Cell Line and Cell Culture

The non-transfected, spontaneously immortalised normal human conjunctival epithelial (IOBA-NHC) cell line was a gift from Dr. Yolanda Diebold at the University of Valladolid, Spain (114). Cells were maintained in a modified medium consisting of 1:1 mixture of Dulbecco's Modified Eagle Medium/Ham's F-12 culture media (Invitrogen, Carlsbad, CA, USA.) with 10% fetal bovine serum (Invitrogen), 2 ng/mL of recombinant human epidermal growth factor (Invitrogen), 1 µg/mL of bovine insulin (Sigma Aldrich), 0.1 µg/mL of cholera toxin, (Sigma Aldrich), 0.5 µg/mL of hydrocortisone and 1× strength of broad spectrum antibiotic-antimycotic (Invitrogen). Prior to the experiment, cells were subcultured, grown to 80% confluence and adapted to Keratinocyte serum-free medium (KSFM) (Invitrogen) media overnight.

The cells were further cultured for 24 h in KFSM at 280, 380 and 480 mOsm for the control, treatments T1 and T2 respectively. The treatment conditions were selected to represent elevated tear osmolarity found in patients with dry eye (115). Five biological replicates were prepared for global metabonomic analysis, and three biological replicates each were prepared for targeted metabonomics, proteomic analysis and Western blot for UDP-*N*-acetylhexosamine

pyrophosphorylase (UAP1) (Figure 3.1). Each hyperosmotic medium was prepared by the addition of 2 M sodium chloride prepared in phosphate-buffered saline (PBS) to commercial KFSM. Osmolarity measurements were performed using a vapour pressure osmometer (Wescor Inc., Logan, UT, USA.). Detailed sample preparation methods are provided in the appendices.

3.2.2. Metabonomic Analysis

Analysis was performed on a Prominence UFLC system (Shimadzu) coupled to an TripleTOF 5600 Q-TOF mass spectrometer (AB Sciex). A T3 C_{18} 2.1 × 100 mm, 3 µm column (Waters Corporation) was used for separation. Each sample was reconstituted in 50 µL of water and 10 µL was injected for each analysis in the positive and negative ionisation polarities. The mobile phase for the positive ionisation polarity was 0.1% formic acid in water (eluent A) and 0.1% formic acid in ACN (eluent B). The linear gradient profile was 2% B from 0 to 2 min, 15% B at 8 min, 50% B at 14 min, 90% B from 18 to 20.5 min, and 2% B from 21 to 27 min. The mobile phase for the negative ionisation polarity was 5 mM ammonium formate and 0.05% formic acid in water (eluent A), and 5 mM ammonium formate and 0.05% formic acid in 9:1 ACN/water (eluent B). The linear gradient profile was 2% B from 0 to 2 min, 15% B at 8 min, 50% B at 14 min, 100% B from 18 to 20.5 min, and 2% B from 21 to 27 min. The flow rate in both chromatographic methods was 0.3 mL/min.

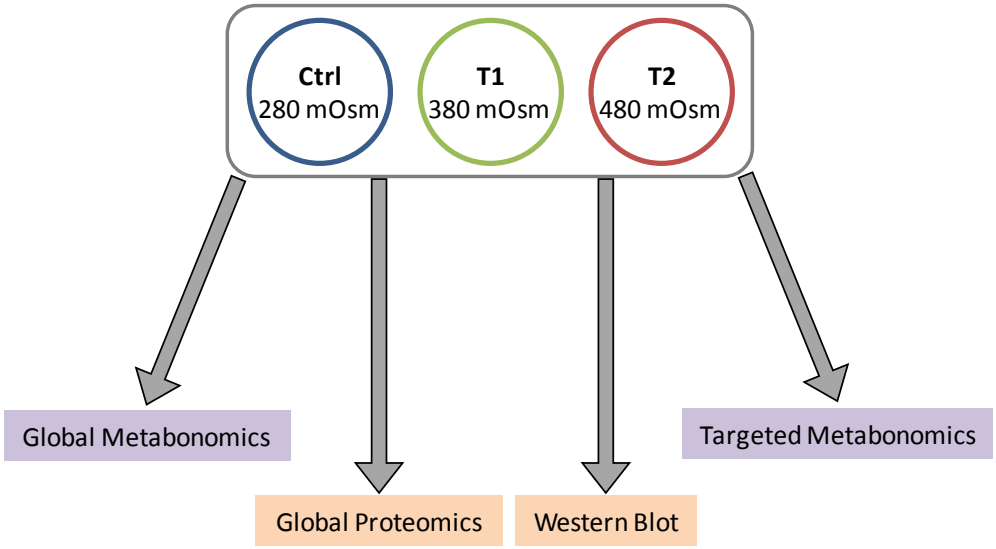


Figure 3.1 Schematic of experiments performed to investigate changes in expression of metabolites and proteins in IOBA-NHC cells under different degrees of hyperosmotic stress.

Acquisition of MS/MS spectra in both ionisation polarities was controlled by the IDA function of Analyst TF. One TOF-MS survey scan (150 ms) was followed by 25 MS/MS scans (50 ms each). The mass range of TOF-MS and MS/MS scans were *m/z* 50–1000 and 25–1000 respectively. The source voltage was 5.0 kV for positive ionisation and 4.0 kV for negative ionisation polarities respectively. The following parameters were applied: dynamic background subtraction, charge monitoring to exclude multiply charged ions and isotopes, and dynamic exclusion of former target ions for 5 s. CE was ramped from 20–30 V in MS/MS scans.

3.2.3. Peak Finding and Alignment of Metabonomic Data

MarkerView (AB Sciex) was used for peak detection and alignment of raw chromatographic data obtained from the respective positive and negative ionisation polarities. Peak detection and alignment parameters are provided in Table 8.1 in the appendices. One peak table comprising the *m/z,* RT and integrated

peak area of all variables present in analysed samples was generated for each ionisation polarity. The following peak classes were excluded from subsequent analysis: (i) peaks attributed to isotopes, (ii) peaks with an average area that were less than three times the average area in blank injections and (iii) peaks with a coefficient of variation (CV) greater than 80% among replicates. The samples in each peak table were normalised to their respective 90^{th} percentile peak area sums.

3.2.4. Statistical Analysis and Metabolite Identification

The Kruskal–Wallis test (a non-parametric analysis of variance) was used to identify *m/z*-RT metabolite peaks with statistically significant median peak areas in at least one group (control, T1 or T2) when compared with other groups. Subsequently, the Mann-Whitney U test was used to identify *m/z*-RT metabolite peaks with integrated peak areas that were significantly different in treatments T1 or T2 as compared with control samples. Metabolite fold-changes between treatment and control groups were calculated using the arithmetic mean. The *m/z* value of each marker peak was queried in METLIN metabolite database (91) to obtain its putative metabolite identification. For this step, the batch search mode was adopted, with mass accuracy set to 12 ppm. The following adducts were applied to the *m/z* values of marker peaks obtained via positive ionisation polarity: $[M+H]^+$, $[M+NH_4]^+$ $[M+Na]^+$, $[M+K]^+$ and $[M-H_2O+H]^+$. Similarly, $[M-H]^-$, $[M-H_2O-H]^-$, $[M+Cl]^-$ and $[M+FA-H]^-$ were applied to *m/z* values obtained via the negative ionisation polarity. Metabolite were identified by retrieval of experimental MS/MS spectra and comparison with the respective entries in METLIN (91) or MassBank (92) mass spectra databases. Unresolved metabolites with the same molecular formula were reported as mixtures.

3.2.5. Targeted Metabonomic Analysis

The levels of carnitine, glycerophosphocholine and phosphocholine were verified in independently-prepared replicates of IOBA-NHC cell lysates. Carnitine, glycerophosphocholine and phosphocholine were purchased from Sigma Aldrich. Separation was performed on a Prominence UFLC system (Shimadzu) with a ZIC HILIC 2.1 × 100 mm, 3.5 µm column (Merck SeQuant). The mobile phase for was 5 mM ammonium formate and 0.05% formic acid in water (eluent A), and 5 mM ammonium formate and 0.05% formic acid in 9:1 ACN/water. The linear gradient profile was 80% B at 0 min, 40% B at 8 min, 20% B from 9 to 11 min, and 80% B from 11 to 15 min. The flow rate was 0.4 mL/min. MRM was performed on a QTrap 6500 ion trap mass spectrometer (AB Sciex). The optimised MRM parameters for each target compound are shown in Table 8.2 in the appendices.

3.2.6. Proteomic Analysis

Samples were trypsin-digested with filter-aided sample preparation (FASP) kits (Protein Discovery, San Diego, CA, USA.) and labelled with isobaric tags for relative and absolute quantification (iTRAQ) reagents (AB Sciex) for proteomic analysis (116). In this study, reagent 114 was added to control samples, reagent 115 was added to T1 samples and reagent 116 was added to T2 samples. After labelling, one control sample, one T1 sample and one T2 sample were pooled together for injection onto the LC-MS/MS system. For each peptide sequence, labelled peptides from different samples give the same *m/z* in TOF-MS but produce reporter ions with different *m/*z in MS/MS. Thus, the ratio of signal intensities for *m/z* 115*:*114 in a MS/MS spectrum of a peptide gives the relative amount of that peptide in treatment T1 compared with the control, and the ratio of signal intensities for *m/z*

116:114 gives the relative amount of that peptide in treatment T2 compared with the control.

Chromatographic separation of peptides was performed on an Ultimate 3000 nano-LC system (Dionex, Thermo Scientific, Sunnyvale, CA, USA). Samples were first loaded onto a 75 µm × 2 cm Acclaim PepMap 100 C_{18} trap column (Dionex, Thermo Scientific) at a flow rate of 5 µL/min. After a 5 min wash with 0.1% formic acid and 2% ACN in water, the trap column was switched to be in-line with a 75 µm × 15 cm Acclaim PepMap RSLC C_{18} analytical column (Dionex, Thermo Scientific). The mobile phase was 0.1% formic acid in water (eluent A) and 0.1% formic acid in 98:2 v/v of ACN/water (eluent B). The linear gradient profile was 5% B at 0 min, 7% B at 5 min, 60% B at 97 min, and 95% B from 100 to 105 min. The flow rate was 300 nL/min.

Data was acquired via the IDA mode using the Analyst TF 1.6 software (AB Sciex). The source voltage was 2.5 kV and the DP was 100 V. The interface heater temperature was 125 °C. The curtain and nebuliser gas flows were 30 and 12 units respectively. The IDA parameters are as follows: 0.25 s of TOF-MS survey scan (*m/z* range: 350–1250) was followed by thirty product ion scans of 75 ms each (*m/z* range: 100–1800). The switching criteria was set to include only ions with charge states of 2 to 5, and an abundance threshold of over 100 counts for fragmentation. Former target ions were excluded for 12 s. CE was controlled by the IDA rolling CE parameter script.

3.2.7. Protein Identification and Data Analysis

ProteinPilot (AB Sciex, version 4.5) was used to analyse MS/MS data output, and MS/MS spectra were searched against the UniProt database (September 2010 release, 40516 proteins in database). Important settings in ProteinPilot were configured as follows: (i) sample type: iTRAQ 4plex (peptide labelled), (ii) Cys alkylation: iodoacetamide, (iii) digestion: trypsin, (iv) instrument: TripleTOF 5600, (v) special factors: none, (vi) ID focus: biological modifications, (vii) search effort: thorough ID and (viii) application of bias correction and background. 95% confidence level was used at the peptide level. False discovery rate (FDR) analysis in the ProteinPilot software was performed and FDR < 1% was set for protein identification. Reverse database search strategy was used to calculate FDR for peptide identification. The Pro Group algorithm in ProteinPilot was used to calculate peak areas of reporter ions. Auto bias correction was applied to eliminate possible pipetting errors in the combination of labelled samples.

Proteins were considered significantly up-regulated if their log-transformed iTRAQ ratios (treatment T1: controls or treatment T2: controls) were above one standard deviation (SD) of the global mean in at least two out of three replicates, and if their fold-changes were greater than 1.4. Proteins were considered significantly down-regulated if their iTRAQ ratios were below one SD of the global mean in at least two out of three replicates, and if their fold-changes were less than 0.71. The average of the global mean iTRAQ ratios obtained in our experiments was 0.99. The average of the global SD of log-transformed iTRAQ ratios was 0.15. The average upper and lower thresholds were 1.4 and 0.71 respectively. Protein fold-changes between control and treatment groups were calculated using the geometric mean in the normal space.

3.2.8. Western Blot for UAP1

Denaturing electrophoresis was performed on lysates and proteins were transferred to nitrocellulose membranes. The membranes were blocked with 5% milk in Tris-buffered saline with Tween 20 before incubation with UAP1 antibody (1:500) (Sigma Aldrich) for 16 h at 4 °C. Protein bands were visualised by using the Supersignal West Pico chemiluminescence kit (Thermo Scientific, Waltham, MA. USA). The blot was also probed with β-actin antibody (1:5000) (Sigma Aldrich) to function as an internal control.

3.3. Results and Discussion: Markers from *In Vitro* Model of Dry Eye Disease

3.3.1. Metabonomic Analysis

There were 1940 *m/z*-RT features from the positive ionisation polarity and 1386 *m/z*-RT features from the negative ionisation polarity. Fifteen dysregulated metabolites were identified by the positive ionisation experiments, while eight were identified by the negative ionisation experiments (Table 3.1, Figure 3.2). Methionine and tyrosine were identified in both the positive and negative ionisation polarities. Several of these dysregulated metabolites were previously reported in human tear samples (117), making them potential biomarkers of conjunctival cell function in dry eye disease. Experimental MS/MS spectra of glycerophosphocholine and UDP-*N*-acetylglucosamine are shown in Figure 3.3 and Figure 3.4 respectively. Data files from global metabonomic analysis have been deposited to MetaboLights (118) with the dataset identifier MTBLS214 (http://www.ebi.ac.uk/metabolights/MTBLS214).

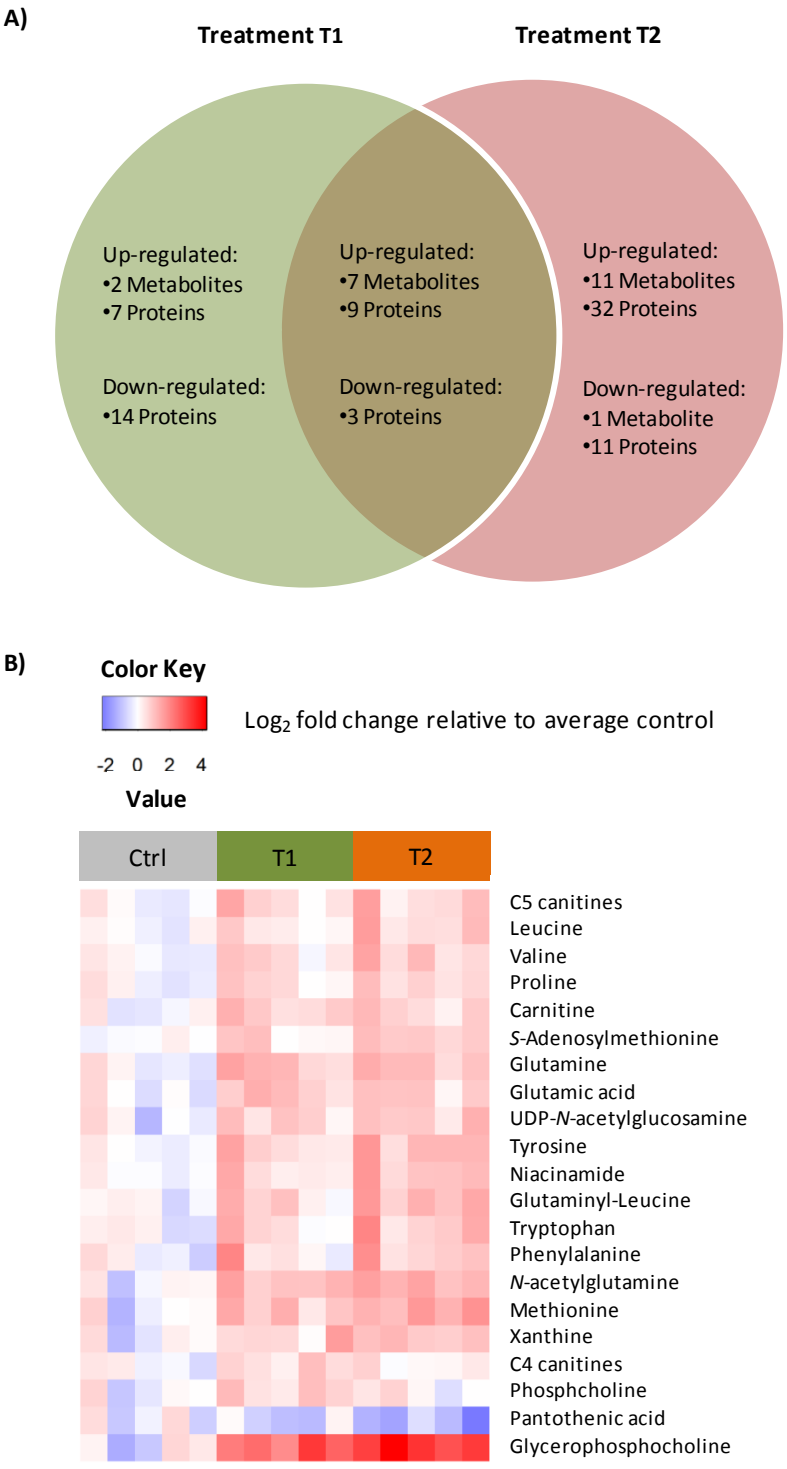


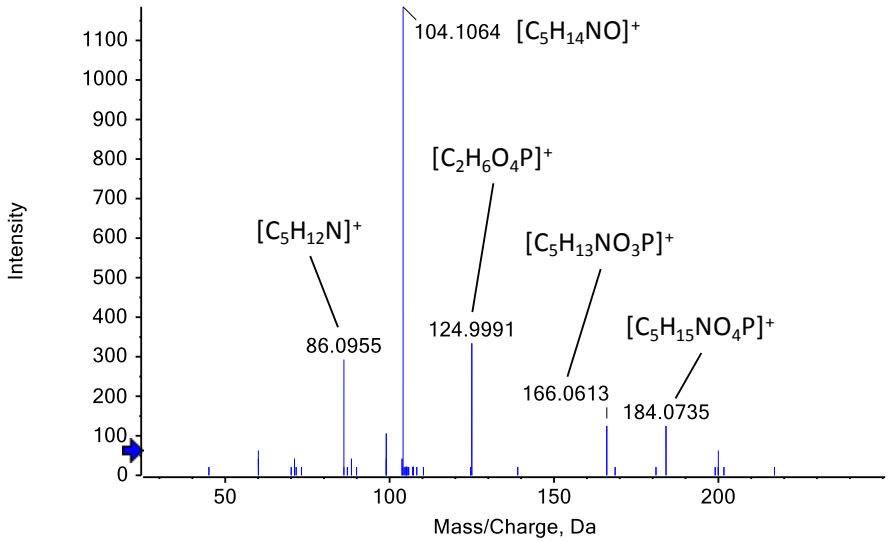
Figure 3.2 A) Summary of the numbers of unique and overlapping dysregulated metabolites and proteins in IOBA-NHC cells under hyperosmotic stress. B) Heatmap of dysregulated metabolites in IOBA-NHC cells under hyperosmotic stress.

Table 3.1 Metabolites in IOBA-NHC cells dysregulated by hyperosmotic stress.

Metabolite	Detected Ion	*m/z*	RT (min)	Fold-change, compared with control	
				T1	T2
Butyryl-carnitine/ Isobutyryl-carnitine	$[M+H]^+$	232.155	5.49	1.59‡	1.24
Carnitine*,†	$[M+H]^+$	162.113	0.89	1.84‡	1.70‡
Glutamic acid†	$[M-H]^-$	146.046	0.93	1.92‡	1.81‡
Glutamine*,†	$[M-H]^-$	145.061	0.86	2.16‡	2.11‡
Glutaminyl-Leucine	$[M+H]^+$	260.161	4.76	1.70	2.43‡
Glycerophosphocholine*,†	$[M+H]^+$	258.109	0.88	6.00‡	11.3‡
Leucine*	$[M+H]^+$	132.103	2.37	1.31	1.94‡
Methionine*,†	$[M-H]^-$	148.043	1.42	2.04‡	1.94‡
	$[M+H]^+$	150.058	1.42	2.05‡	2.77‡
N-acetylglutamine	$[M-H]^-$	187.072	1.04	2.22‡	2.51‡
Niacinamide†	$[M+H]^+$	119.048	2.27	1.60	2.23‡
Panthothenic acid*	$[M-H]^-$	218.103	5.14	0.74	0.42‡
Phenylalanine*,†	$[M+H]^+$	166.086	4.49	1.74	2.10‡
Phosphocholine†	$[M+H]^+$	184.074	0.86	1.74‡	1.17
Proline	$[M+H]^+$	116.071	0.96	1.49	1.65‡
S-Adenosylmethionine†	$[M+H]^+$	399.145	0.85	1.46	1.86‡
Tryptophan*	$[M+H]^+$	205.097	6.52	1.58	2.31‡
Tyrosine*	$[M-H]^-$	180.065	1.71	1.79‡	1.85‡
	$[M+H]^+$	182.082	2.29	1.75‡	2.37‡
UDP-*N*-acetylglucosamine*,†	$[M-H]^-$	606.076	0.96	1.69	1.91‡
2-Methylbutyroylcarnitine/ Isovalerylcarnitine/ Valeryl-carnitine	$[M+H]^+$	246.171	7.62	1.69	1.89‡
Valine*,†	$[M+H]^+$	118.086	1.33	1.55	1.91‡
Xanthine*,†	$[M-H]^-$	151.026	2.00	1.77	2.02‡

* Verified with standard
† Previously detected in human tear samples
‡ Significant between treatment and control group

A)



B)

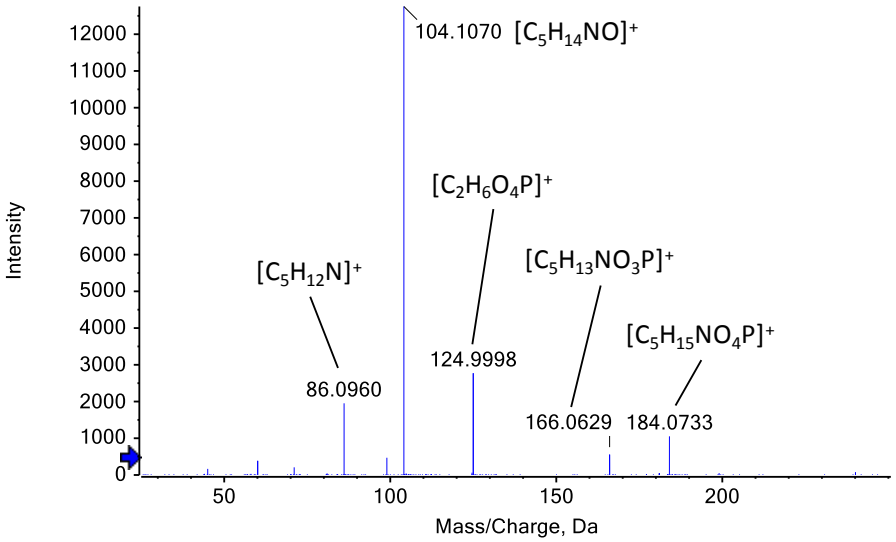
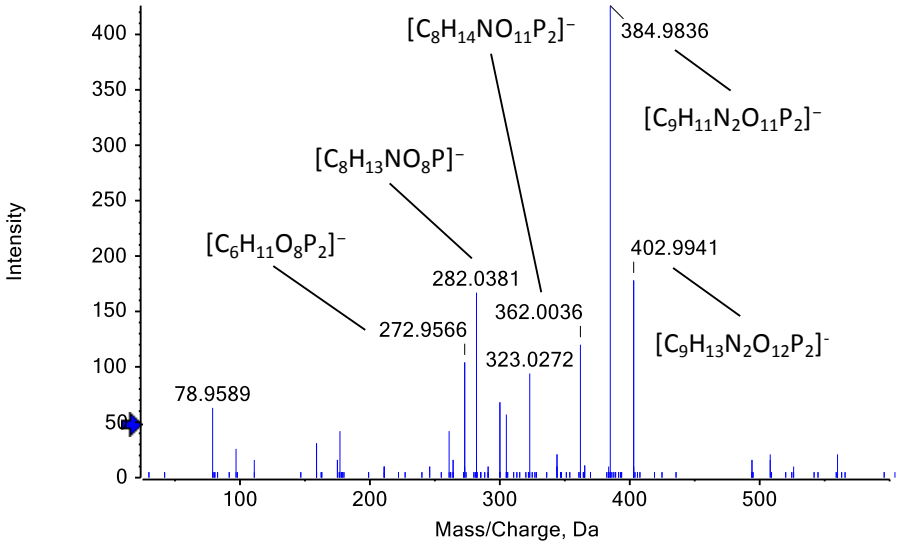


Figure 3.3 A) MS/MS spectra of glycerophosphocholine derived from metabonomic analysis of IOBA-NHC cell lysate and B) verified with standard analysed under same LC-MS/MS conditions.

A)



B)

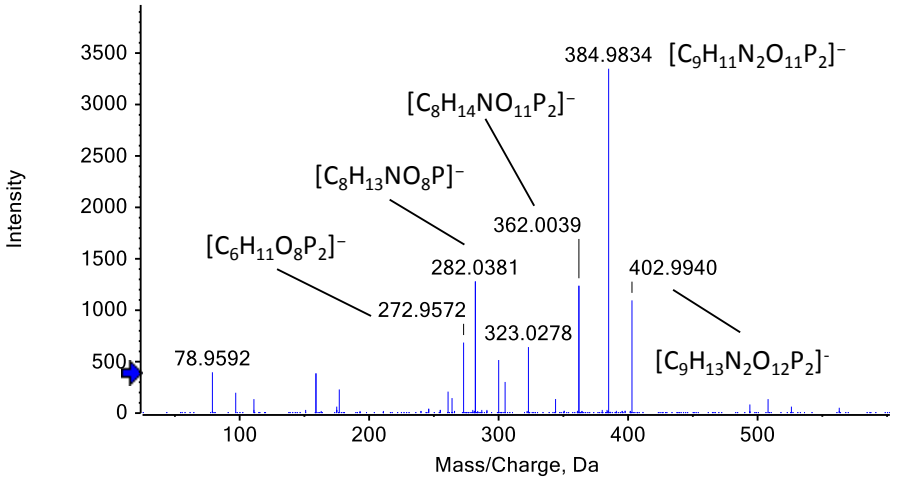


Figure 3.4 A) MS/MS spectra of UDP-*N*-acetylglucosamine derived from metabonomic analysis of IOBA-NHC cell lysate and B) verified with standard analysed under same LC-MS/MS conditions.

3.3.2. Osmoprotectant Metabolites

Several groups have investigated the effect of small molecule osmoprotectants using *in vitro* and animal models of dry eye. Corrales *et al*. found that supplementation of carnitine or erythritol in culture media reduced the levels of activated JNK and p38 MAPK in human corneal epithelial cells under hyperosmotic stress (119). Garett *et al.* found the addition of betaine to culture media reduced apoptosis in human corneal limbal epithelial cells (120). More recently, Hua *et al.* foundy that carnitine, betaine and erythritol suppressed the expression of pro-inflammatory cytokines in human corneal epithelial cells cultured in hyperosmotic media (121). Chen *et al.* established that carnitine, betaine and erythritol reduced the expression of pro-inflammatory cytokines in the conjunctiva and reduced apoptosis in the corneal tissues in a mouse model of dry eye (122).

In this study, carnitine levels in IOBA-NHC cells were increased under treatments T1 and T2, in agreement with its function as an osmoprotectant. In addition, the levels of short-chain carnitines were found to be increased under hyperosmotic conditions although dose-dependency with increasing osmolarity was not observed. While betaine was detected in the positive ionisation polarity (*m/z* 118.086, RT 0.88 min) in metabonomic analysis, it did not show any significant increase under either hyperosmotic stress treatment ($p>0.05$). Erythritol was not detected in IOBA-NHC cells in our study. Our results show that carnitine is a preferred endogenous osmoprotectant over betaine.

Among the dysregulated metabolites identified from global metabonomic analysis, glycerophosphocholine showed the largest fold-change and a dose-dependent relationship with increasing osmotic stress (Table 3.1, Figure 3.2). An increase in

phosphocholine, a key intermediate in glycerophosphocholine synthesis was also observed under treatment T1 but not treatment T2 (Figure 3.5). This suggests that an overall activation of this pathway under hyperosmotic stress and a depletion of intermediates at high osmotic stress. Although glycerophosphocholine is a well-studied osmolyte in renal medullary cells (123, 124) and has been detected in rabbit corneas (16), this study is the first to report its involvement the osmotic stress response in ocular surface cells. The dose-dependent elevation in glycerophosphocholine levels suggests that it is more readily-accumulated than other osmoprotectants in IOBA-NHC cells, and its accumulation is a more sensitive response to hyperosmotic stress.

To investigate if glycerophosphocholine is the preferred endogenous osmolyte over carnitine, these metabolites from three independent biological replicates were quantified using MRM (Figure 3.1). Carnitine levels showed a 1.4 fold increase under treatment T1 and 2.1 fold increase under T2 (Table 3.2). Glycerophosphocholine showed a 2.1 fold increase under treatment T1 and 5.8 fold increase under T2 (Table 3.2), showing dose-dependent up-regulation of glycerophosphocholine with increasing osmotic stress. Collectively, these results suggest that glycerophosphocholine is the preferred endogenous osmolyte in IOBA-NHC cells. It would be pertinent to investigate the efficacy of glycerophosphocholine in reducing inflammation and apoptosis in *in vitro* and animal models when supplemented to culture media and eye drops respectively.

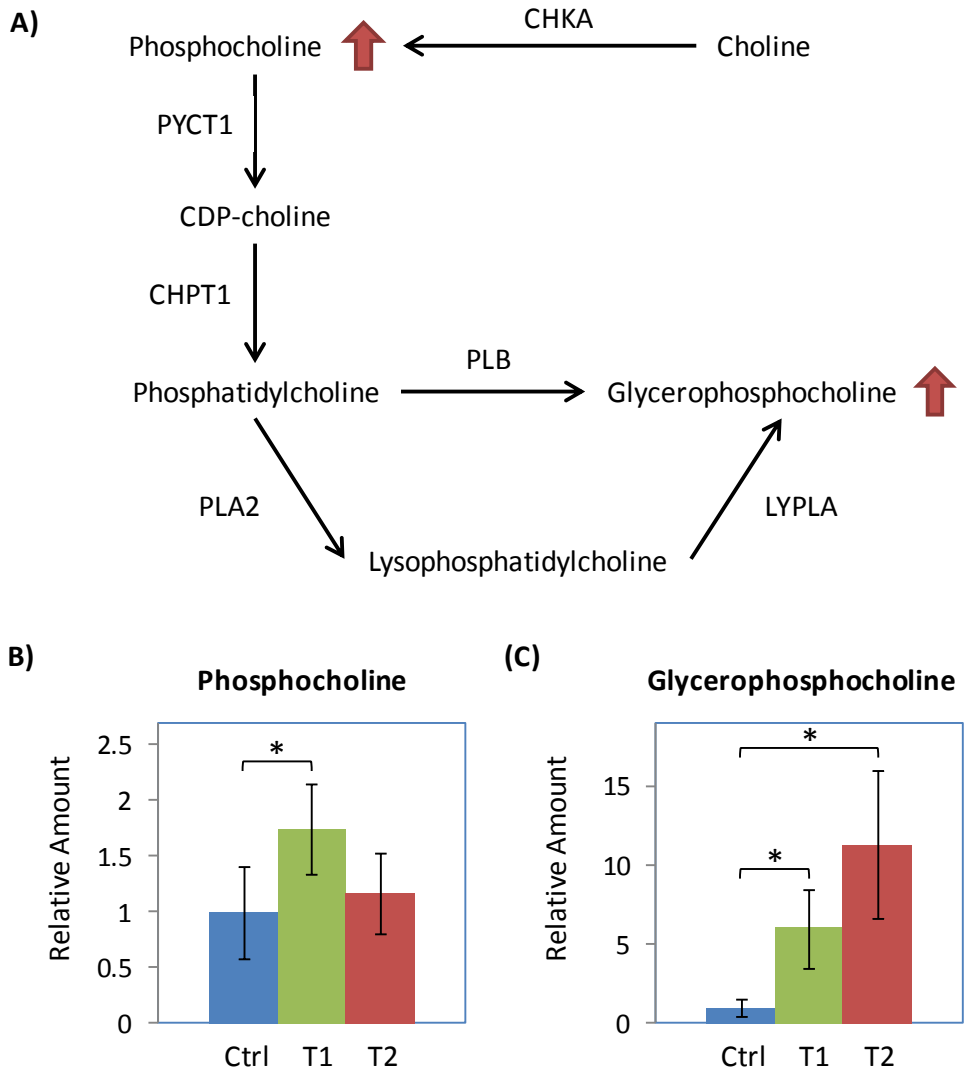


Figure 3.5 A) Synthesis pathway of glycerophosphocholine, CHKA: choline kinase alpha, PYCT1: choline-phosphate cytidylyltransferase, CHPT: cholinephosphotransferase 1, PLB: phospholipase B1, PLA2: phospholipase A2:, LYPLA: lysophospholipase. B) Relative amount of phosphocholine in control, T1 and T2 cell lysates. C) Relative amount of glycerophosphocholine in control, T1 and T2 cell lysates.

Table 3.2 MRM-validated fold-changes of carnitine, glycerophosphocholine and phosphocholine levels in IOBA-NHC cells under hyperosmotic stress.

Metabolite	Fold-change, compared with control	
	T1	T2
Carnitine	1.43	2.13
Glycerophosphocholine	2.68	5.83

3.3.3. Proteomic Analysis

A total of 2373 proteins were identified from proteomic analyses, with 1483 common proteins among the three replicates. Only proteins with at least two unique peptides were considered for statistical analysis. Compared with controls, there were 26 proteins with increased levels and 20 proteins with decreased levels under treatment T1 (380 mOsm), and 45 proteins with increased levels and 20 proteins with decreased levels under treatment T2 (480 mOsm). Between the two treatments, there were 15 proteins in common with increased levels and 7 proteins in common with decreased levels (Figure 3.2). The list of dysregulated proteins is shown in Table 3.3 and a heatmap of these proteins are shown in Figure 3.6.

Data from iTRAQ-based proteomics have been deposited to the ProteomeXchange Consortium (125) via the PRIDE partner repository (126) with the dataset identifier PXD002661 (http://www.ebi.ac.uk/pride/archive/projects/PXD002661).

Table 3.3 Proteins in IOBA-NHC cells dysregulated under hyperosmotic stress.

Gene Symbol	Accession (UniProt/SwissProt KB)	Protein Name	Fold-change, compared with control	
			T1	T2
AKR1B1	P15121	Aldose reductase	1.47*	2.06*
ANXA1	P04083	Annexin A1	1.27	1.85*
ATP1A1	P05023	Sodium/potassium-transporting ATPase subunit alpha-1	1.25	1.60*
ATP5C1	P36542	ATP synthase subunit gamma, mitochondrial	1.31	1.46*
BZW1	Q7L1Q6	Basic leucine zipper and W2 domain-containing protein 1	1.71*	1.95*
CANX	P27824	Calnexin	1.26	1.53*
CCT7	Q99832	T-complex protein 1 subunit eta	0.64*	0.82
CPNE3	O75131	Copine-3	0.54*	0.45
CPS1	P31327	Carbamoyl-phosphate synthase [ammonia], mitochondrial	1.30	1.49*
DDOST	P39656	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	1.57*	1.56*
DDX3X	O00571	ATP-dependent RNA helicase DDX3X	1.36	1.56*
DDX5	P17844	Probable ATP-dependent RNA helicase DDX5	0.91	1.67*
DNAJA2	O60884	DnaJ homolog subfamily A member 2	0.67*	0.99
DPH5	Q9H2P9	Diphthine synthase	0.48*	0.90
DYNC1H1	Q14204	Cytoplasmic dynein 1 heavy chain 1	0.70*	0.79
EEF1A2	Q05639	Elongation factor 1-alpha 2	0.94	0.65*
EEF2	P13639	Elongation factor 2	1.14	1.54*
EIF3D	O15371	Eukaryotic translation initiation factor 3 subunit D	0.68*	0.68

Gene Symbol	Accession (UniProt/SwissProt KB)	Protein Name	Fold-change, compared with control	
			T1	T2
EIF4A1	P60842	Eukaryotic initiation factor 4A-I	1.28	1.67*
EIF4A3	P38919	Eukaryotic initiation factor 4A-III	0.99	1.66*
EIF5	P55010	Eukaryotic translation initiation factor 5	1.05	1.65*
EIF6	P56537	Eukaryotic translation initiation factor 6	0.64*	0.98
EPRS	P07814	Bifunctional aminoacyl-tRNA synthetase	0.70*	0.79
ERP29	P30040	Endoplasmic reticulum resident protein 29	1.31	1.41*
GNB1	P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	0.19*	0.72
GRWD1	Q9BQ67	Glutamate-rich WD repeat-containing protein 1	1.79*	0.77
HIST1H4A	P62805	Histone H4	1.58*	1.75*
HSPA5	P11021	78 kDa glucose-regulated protein (Heat shock 70 kDa protein 5)	1.13	1.58*
HYOU1	Q9Y4L1	Hypoxia up-regulated protein 1	0.79	0.60*
KRT10	P13645	Keratin, type I cytoskeletal 10	0.87	1.56*
KRT7	P08729	Keratin, type II cytoskeletal 7	0.65*	0.72
KRT8	P05787	Keratin, type II cytoskeletal 8	1.23	1.56*
LCP1	P13796	Plastin-2	0.76	0.63*
LDHA	P00338	L-lactate dehydrogenase A chain	1.31	1.60*
LSM7	Q9UK45	U6 snRNA-associated Sm-like protein	1.21	2.09*
MAP2K3	P46734	Dual specificity mitogen-activated protein kinase kinase 3	0.87	1.62*
MYO1G	B0I1T2	Myosin-Ig	0.65*	0.87
NAP1L1	P55209	Nucleosome assembly protein 1-like 1	0.90	1.58*
NCL	P19338	Nucleolin	1.37	1.72*

Gene Symbol	Accession (UniProt/SwissProt KB)	Protein Name	Fold-change, compared with control	
			T1	T2
NSUN2	Q08J23	tRNA (cytosine-5-)-methyltransferase	0.52*	0.98
NUP155	O75694	Nuclear pore complex protein	0.65*	0.67*
OGDH	Q02218	2-oxoglutarate dehydrogenase, mitochondrial	1.22	1.48*
PDHB	P11177	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	1.49*	0.96
PKM2	P14618	Pyruvate kinase isozymes M1/M2	1.13	1.53*
PLEC	Q15149	Plectin	1.42*	2.17*
PPIB	P23284	Peptidyl-prolyl cis-trans isomerase B	1.19	1.41*
PRDX1	Q06830	Peroxiredoxin-1	1.23	1.77*
PRDX2	P32119	Peroxiredoxin-2	0.82	0.64*
PRPS1	P60891	Ribose-phosphate pyrophosphokinase 1	0.71	0.66*
PSMD1	Q99460	26S proteasome non-ATPase regulatory subunit 1	0.65*	0.66*
PSMD2	Q13200	26S proteasome non-ATPase regulatory subunit 2	0.68*	0.82
PTGS2	P35354	Prostaglandin G/H synthase 2	1.75*	3.35*
PTMA	P06454	Prothymosin alpha	0.68	1.56*
PXN	P49023	Paxillin	0.91	0.70*
PYGB	P11216	Glycogen phosphorylase, brain form	1.45	1.46*
RHOC	P08134	Rho-related GTP-binding protein RhoC	1.52*	1.45
RPL21	P46778	60S ribosomal protein L21	0.98	1.61*
RPN2	P04844	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2	1.64*	1.75*
SEC23A	Q15436	Protein transport protein Sec23A	1.62*	1.09

Gene Symbol	Accession (UniProt/SwissProt KB)	Protein Name	Fold-change, compared with control	
			T1	T2
SERBP1	Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	1.45*	1.43
SERPINB3	P29508	Serpin B3	1.02	0.60*
SLPI	P03973	Antileukoproteinase	1.16	1.74*
STOM	P27105	Erythrocyte band 7 integral membrane protein	1.50*	1.14
TGM2	P21980	Protein-glutamine gamma-glutamyltransferase 2	0.68*	0.57*
TIMP1	P01033	Metalloproteinase inhibitor 1	0.88	1.53*
TOMM34	Q15785	Mitochondrial import receptor subunit T	0.44*	0.79
TOMM6	Q96B49	Mitochondrial import receptor subunit T	1.43*	1.03
TPBG	Q13641	Trophoblast glycoprotein	1.34	1.48*
TUBB2C	P68371	Tubulin beta-2C chain	1.17	1.65*
UAP1	Q16222	UDP-N-acetylhexosamine pyrophosphorylase	1.79*	2.87*
UBXN1	Q04323	UBX domain-containing protein 1	1.00	0.70*
VDAC1	P21796	Voltage-dependent anion-selective channel protein 1	1.49*	1.48*
VIM	P08670	Vimentin	1.32	1.54*
WARS	P23381	Tryptophanyl-tRNA synthetase, cytoplasmic	0.90	0.59*
XRCC5	P13010	X-ray repair cross-complementing protein 5	0.77	0.71*
XRN2	Q9H0D6	5'-3' exoribonuclease 2	1.55	0.59*

*significant change from the global mean in at least two out of three iTRAQ analysis

Figure 3.6 Heatmap of dysregulated proteins in IOBA-NHC cells under hyperosmotic stress.

3.3.4. Osmotic Stress Response

The heat shock protein (HSP) response is a well-studied adaptation to cellular stress (127). While fourteen HSPs were detected in IOBA-NHC lysates, only HSPA5 showed increased levels under treatment T2. As the cultures were treated for 24 h, it was not possible to detect rapid and transient changes in HSP levels in response to hyperosmotic stress, as Li *et al*. had demonstrated with HEK293 cultures (128). Li *et al*. had also found HSPA5 levels to be persistently elevated when HEK293 cells were cultured in hyperosmotic media for up to two passages (128). These results suggest that HSPA5 up-regulation could be a response to chronic hyperosmotic stress in ocular surface tissue. This study also identified increased levels of intermediate filament proteins, such as keratin 8 and vimentin, that are suggested to play cytoprotective roles similar to the classic heat shock response (129).

3.3.5. Hexosamine Pathway

In this study, a significant increase in UDP-*N*-acetylglucosamine was observed under T2 (Table 3.1, Figure 3.7). UAP1, the enzyme responsible for UDP-*N*-acetylglucosamine synthesis, was increased under T1 and T2 (Table 3.1, Figure 3.7). UDP-*N*-acetylglucosamine is a key intermediate between the hexosamine and cell signalling pathway as it provides the substrate for *O*-linked β-*N*-acetylglucosamine glycosylation (*O*-GlcNAcylation) of serine and threonine residues. This post-translational modification is critical in modulating protein function, degradation and localisation; and is suggested to be a rapid survival response to cellular stress. (130)., demonstrating concordance between findings gleaned from both metabonomic and proteomic analysis. The increase in UAP1 levels under hyperosmotic stress was confirmed by Western blot (Figure 3.7).

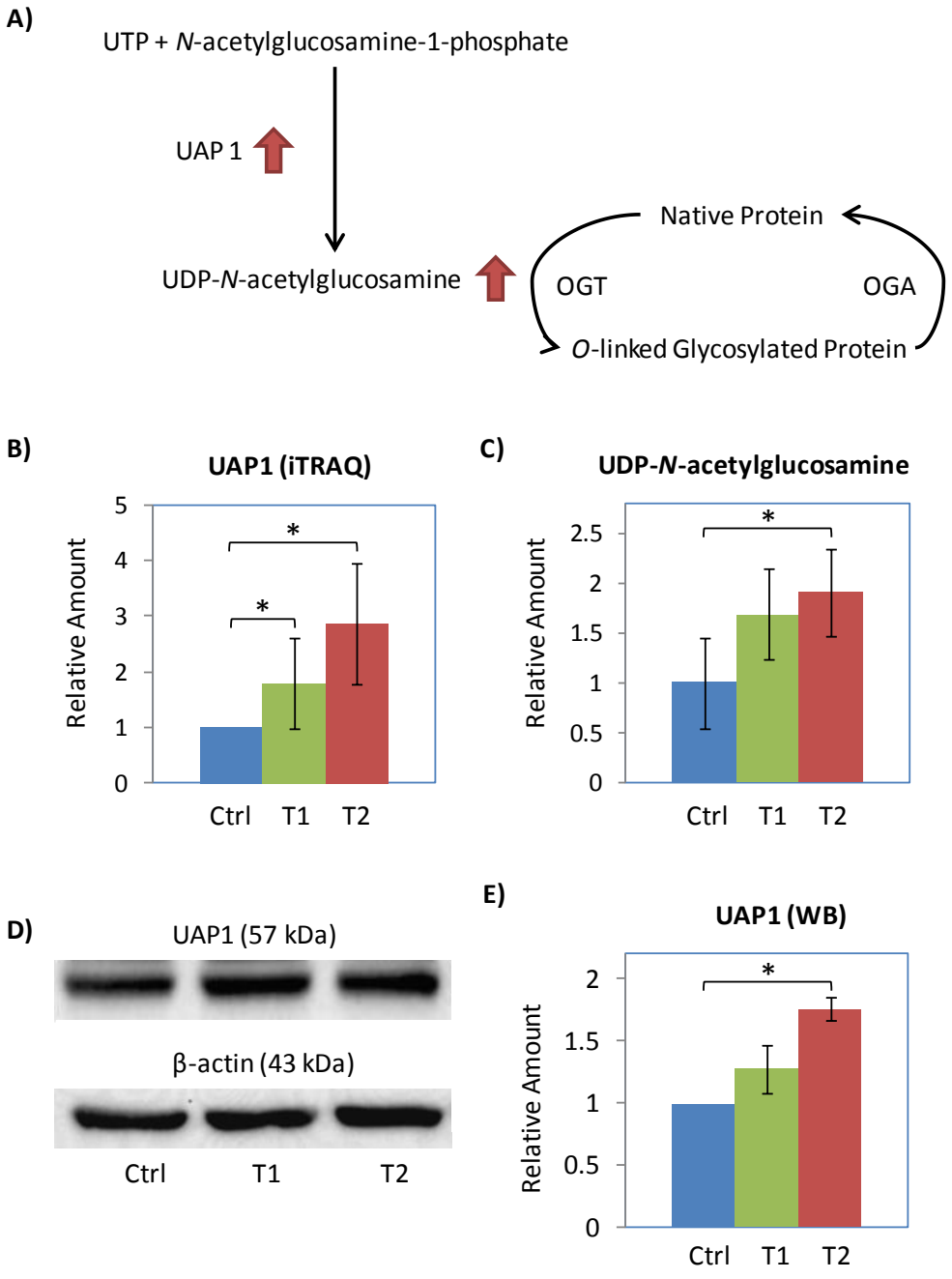


Figure 3.7 A) Synthesis pathway of glycerophosphocholine, OGT: *O*-GlcNAc transferase, OGA: *O*-GlcNAc-ase. B) Relative amount of UAP1 in control, T1 and T2. C) Relative amount of UDP-*N*-acetylglucosamine in control, T1 and T2 cell lysates. D) Representative Western blot of UAP1 and E) UAP1 quantified by densitometry.

Herzog *et al*. recently showed that increasing the levels of *O*-GlcNAcylation in mesothelial cells improved cell viability when the cultures were treated were hyperosmotic peritoneal dialysis fluid (131). Targeting the levels of *O*-GlcNAcylation in ocular surface tissues may be a novel therapy to potentiate cell survival in dry eye disease.

3.3.6. Summary for Metabonomic and Proteomic Analysis of an *In Vitro* Model of Dry Eye Disease

Using global metabonomic analysis, this study identified dose-dependent increases in glycerophosphocholine levels in IOBA-NHC cells under hyperosmotic stress. This trend was verified by targeted metabonomic analysis, showing that glycerophosphocholine is the preferred osmoprotectant over other endogenous metabolites such as carnitine. In addition, this study also indentified activation of the hexosamine pathway from increased levels of UDP-*N*-acetylglucosamine and UAP1 under hyperosmotic stress, identified using global metabonomic analysis and iTRAQ-based proteomic analysis respectively.

4. Plasma Metabotyping of Diabetic Retinopathy

4.1. Introduction to Diabetic Retinopathy

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and the leading cause of visual impairment in working-age adults worldwide (55, 56). The disease is characterised by retinal haemorrhage, aneurysms, neo-vascularisation of the retina. The early stage of DR is asymptomatic (132). The global prevalence of diabetes is rising and the number of diabetic persons is projected to increase by 54% between 2010 and 2030 (133). The public health burden of diabetes and its complications would thus increase correspondingly. In Singapore, the prevalence of diabetes is 23% in the Malay ethnic group and 33% in the Indian (South-Asian) ethnic group (134, 135). The prevalence of DR among diabetic persons in the Malay and Indian ethnic groups in Singapore are 35% and 30% respectively (134, 135).

There is no pharmacological treatment available for the early stages of DR. The major risk factors of DR are poor glycaemic control, hypertension, dyslipidaemia, and the duration of diabetes (136, 137). The effects of risk factor-management on DR prevention and control vary among study populations (56, 138). The results from randomised controlled trials show that a significant proportion of patients assigned to intensive control of metabolic risk factors continue to develop diabetic complications (139, 140). There is increasing evidence to suggest that metabolic memory is responsible for this observation. The term metabolic memory refers to the persistent epigenetic modifications caused by early exposure to hyperglycaemia that in turn predispose individuals to development of diabetic

complications even after good glycaemic control is achieved (141). More recently, Zavrelova *et al*. found that the progression of DR in a treated cohort occurs in distinct patterns, ranging from a stable condition that is free from DR to rapid progression to the proliferative DR within six years (142). The mechanisms responsible for these patterns in disease progression are unclear, and the contribution of the duration of diabetes towards these patterns is also not known (143). The discovery of biomarkers that characterise the different developmental phenotypes of DR thus become important as these biomarkers may provide insight on pathogenic pathways that are currently unknown. With an increase in the prevalence of diabetes (133), there is an interest to stratify patients into low and high risk groups of DR progression, and optimise resources for DR screening.

Recent metabonomic studies have uncovered plasma and sera metabolic signatures associated with, or predictive of impaired glucose tolerance and diabetes (57-66). Furthermore, DR is a complex disease where findings from genome-wide-association studies have not been conclusive (67). We postulate that a metabolic signature of DR exists, and can be resolved from that of diabetes alone. Barba *et al.* identified metabolite markers of DR in the vitreous humour (21). However, the invasive nature of vitreous sampling limits study replication and the translational potential of any biomarkers identified from vitreous fluid. On the other hand, plasma or sera have remained the biofluid of choice in metabonomic studies. With these biological matrices, several research groups have reported metabolite markers of global (22, 144) and targeted (145-147) metabonomic analysis.

These studies however, did not account for confounders and co-morbidities such as medication use and kidney disease. In this chapter, we sought to investigate if the plasma metabotype of DR is resolvable from the metabolic perturbations associated with underlying diabetes, and to identify metabolite markers of DR.

4.2. Methods in Profiling Plasma Samples

4.2.1. Sample Selection

Samples were selected from banked plasma collected as part of the Singapore Indian Eye Study (68). The study was approved by the SingHealth Institutional Review Board, conducted in accordance to the Declaration of Helsinki, and written informed consent was obtained from all participants. Forty samples were selected from diabetic participants with moderate non-proliferative DR (level 43 on the early treatment of diabetic retinopathy scale) (148) or worse in at least one eye. Forty control samples were selected from retinopathy-free diabetic participants with comparable duration of diabetes.

4.2.2. GC-MS Metabonomic Profiling

Plasma samples (100 µL aliquots in microcentrifuge tubes) were thawed to room temperature (25 °C). 20 µL of D-27 myristic acid in methanol (200 µg/mL) and 300 µL of methanol were added to each sample. The tubes was mixed for 15 min at room temperature and centrifuged for 10 min at 16,000 g (4 °C). 300 µL of each supernatant was transferred carefully to a glass tube and dried under nitrogen gas. The dried samples were re-suspended in 100 µL of toluene, vortexed-mixed for 10 s and dried again under nitrogen. A two-step derivatisation method was employed for chemical derivatisation of metabolites. Each sample was incubated with 50 µL

of 2% methoxyamine chloride in pyridine (Pierce Biotechnology, Thermo Scientific) for 1.5 h at 60 °C. Next, 50 µL of *N*-methyl-*N*-trifluoroacetamide (Thermo Scientific) was added and the samples were incubated again at 60 °C for 1 h. 80 µL of derivatised samples were transferred to silanised glass vials for GC-MS analysis. A maximum of thirty-six samples were prepared each day to ensure GC-MS data acquisition was performed within 48 h post-derivatisation. Ten aliquots of pooled human plasma collected with K_2EDTA tubes (Innovative Research Inc, batch 052511), three aliquots of distilled water and 10 mM glucose in distilled water were also subject to the same sample pre-treatment procedure as the plasma samples.

GC-MS analysis was performed with an Agilent 6890N gas chromatograph (Agilent Technologies, Santa Clara, CA, USA) coupled to Leco Pegasus III (4D) GC × GC-MS TOF mass analyser operating in GC-MS mode (Leco Corporation, St. Joseph, MI, USA). Sample injection was performed by a CTC PAL autosampler (CTC Analytics AG, Zwingen, Switzerland). The injection temperature was 250 °C and injection volume was 1 µL, with a spilt ratio of 1:5. A capillary column with DB-1 stationary phase of 21.5 m × 0.25 mm, 0.25 µm film thickness (Agilent Technologies) was used with a constant helium carrier gas flow of 1.5 mL/min. The temperature gradient was held at 70 °C for 2 min, increased at a rate of 10 °C/min to 310 °C, and held at 310 °C for 6 min. The transfer line and ion source temperatures were 280 and 250 °C respectively. Mass spectra were acquired after a delay of 240 s post-injection, from *m*/z 45 to 600 at 20 Hz using ionisation energy of 70 eV and a detector voltage of 1650 V. Blanks and glucose standards were injected at the beginning, in the middle and at the end of the sample batch. The ten pooled

plasma aliquots were injected before the start of the first study sample and after every ten clinical samples to act as quality controls (QCs).

4.2.3. Data Processing

Deconvolution and peak finding were performed with ChromaTOF (version 4.44, Leco Corporation) using these parameters: (i) a S/N threshold of 150, (ii) a minimum peak width of 2.5 s and (iii) a minimum of two apexing masses. Instrument performance was assessed using the peak area of D-27 myristic acid.

Peak alignment was performed with the "calibration" function in ChromaTOF (149). In this method, a reference table was first built using data obtained from plasma samples of participants who had declared no medication usage. There were three such participants (2 DR, 1 non-DR) among the selected samples. An additional six samples were selected from diabetic participants in the Singapore Indian Eye Study who had declared no medication usage for the purpose of building the reference table. Each peak was only included in the reference table for statistical analysis if (i) it was detected in at least seven out of nine medication-free samples (78%) and (ii) its average integrated peak area was at least five times greater than that of the glucose standards and blanks. This strategy was to ensure that the peaks in the reference table included endogenous metabolites as represented in the medication-free samples and excluded xenobiotics such as drugs and their associated metabolites. Peaks derived from glucose were saturated in most samples and were thus excluded from the reference table. Furthermore, plasma glucose levels were not significantly different between DR and non-DR participants in this study.

Data from the analysed plasma samples representing DR and non-DR phenotypes in the absence and presence of kidney disease were then aligned to the reference table. Peaks with missing values in more than 20% of the samples were discarded. For those with missing values in less than 20% of the samples, the missing values were filled with half of the lowest value obtained for that peak across all samples. The raw peak area was normalised to the total peak area of all peaks within the 90th percentile of peak area distribution in each sample, using the median sample as the denominator in the calculation of normalisation factors. Normalised data was then log transformed (base 2). Peaks with CV greater than 30% in the QC samples were not considered for further analysis.

4.2.4. Statistical Analysis

PCA with unit variance scaling was performed to evaluate the presence of dominating trends in the peak data. Mann-Whitney U tests were performed to identify metabolite markers of DR (*p* less than 0.05, fold change greater than 1.2 or less than 0.8). FDR correcting for multiple testing was calculated with the Benjamini-Hochberg method (150). The areas under curve (AUC) in receiver operating characteristic (ROC) analysis were calculated to evaluate the discriminating power of the metabolite markers (151). Logistic regression models were calculated to evaluate the association of metabolite peak areas with the presence of DR. PCA was performed with SIMCA (Umetrics). Mann-Whitney U tests, ROC analyses and logistic regression were performed in the R environment.

4.2.5. Metabolite Identification

Only metabolite markers with a CV less than 30% in the QC samples were identified. Experimentally-obtained spectra were searched against entries in the

National Institutes of Science and Technology (NIST) library (built-in with ChromaTOF) for putative identifications of metabolites. Metabolite identities were confirmed if both the forward and reverse similarity scores were above 800 and the Kovats retention index was within 2% from database entries. Peaks that could not be identified in the NIST library were exported to text spectra and searched against the Golm Metabolome Database (152). Ambiguous identifications were resolved by injection of derivatised analytical standards. For metabolites forming multiple derivatisation peaks, the peaks with lower CV in QC samples were reported. Peaks with unresolvable identifications were reported as mixtures of compounds.

4.2.6. Pathway Analysis

Pathway analysis was performed with MetaboAnalyst (153). Pathways were considered to be significantly enriched in the dataset if their *p*-value was less than 0.05, and their FDR less than 0.2.

4.3. Results and Discussion: Findings from Metabotyping of Diabetic Retinopathy

4.3.1. Sample Characteristics

The clinicopathological characteristics of the selected participants are shown in Table 4.1. Sample groups were comparable for most metabolic and clinical characteristics, such as BMI, total cholesterol, LDL and HDL cholesterol. Participants with DR had greater percentage of HbA1c than their non-DR counterparts, mirroring typical epidemiological findings (135, 136). While increased systolic BP was identified as strong risk factor for DR in the main

Singapore Indian Eye Study, the difference in systolic BP between the DR and non-DR groups in this study was not statistically significant (p=0.086). When participants were further stratified by HbA1c thresholds, clinicopathological variables between DR and non-DR groups remained comparable (p> 0.05, Table 8.3 and Table 8.4 in the appendices).

Table 4.1 Clinical characteristics of participants selected for metabonomic profiling.

Clinical Characteristics*	DR	Non-DR	*P*-value
N	40	40	-
Gender	20 M	23 M	0.501†
Age (years)	59 (53–66)	62 (52–69)	0.846‡
Diastolic BP (mm Hg)	74 (68–81)	77 (68–83)	0.516‡
Systolic BP (mm Hg)	146 (132–155)	133 (122–146)	0.086‡
BMI (kg/m^2)	24.5 (23.0–28.1)	28.6 (23.4–30.3)	0.070‡
Serum creatinine (μM)‡	84 (71–134)	80 (58–117)	0.092‡
Blood glucose, random (mM)	10.4 (6.8–14.6)	8.3 (6.9–11.2)	0.109‡
Total cholesterol (mM)	4.50 (3.99–5.39)	4.38 (3.66–5.25)	0.320‡
LDL cholesterol (mM)	2.84 (2.17–3.49)	2.59 (2.08–3.27)	0.332‡
HDL cholesterol (mM)	0.98 (0.78–1.23)	1.00 (0.84–1.23)	0.486‡
Triglycerides (mM)	1.79 (1.39–2.59)	1.35 (0.94–2.19)	0.076‡
Blood HbA1c (%)‡	8.2 (7.4–10.2)	7.4 (6.5–8.0)	0.001‡
Duration of diabetes (years)	12 (7–21)	11 (5–19)	0.374‡
Urine ACR (μg/mg creatinine)	46.0 (16.8–161.6)	25.7 (13.2–56.4)	0.053‡
eGFR (mL/min/1.73 m^2)	68.7 (44.6–95.9)	87.5 (50.7–100.9)	0.141‡
Microalbuminuria	21	14	0.115†
Macroalbuminuria	5	2	0.235†
Chronic Kidney Disease§	16	11	0.237†
On Diabetes Medication	34	32	0.556†

* Values indicate median with interquartile range, where applicable
† Pearson chi-square test
‡ Mann-Whitney U test
§ Defined as eGFR below 60 mL/min/1.73 m^2

4.3.2. Data Table and Chemometric Analysis

There were 263 peaks in the reference table constructed from medication-free samples and 258 were present in at least 80% of the samples. The CV of D-27 myristic acid, the internal standard was 7.9% and 193 peaks had a CV of less than 30% in the QC samples. The PCA model constructed from aligned peak data was optimised at 5 PCs, with R^2 and Q^2 values of 0.48 and 0.27 respectively. Most samples from participants with chronic kidney disease are located towards the positive scores along the first PC (Figure 4.1A). When labelled by HbA1c values, the majority of samples with greater HbA1c levels are found on the positive axis of the second PC (Figure 4.1B). These clustering trends show that chronic kidney disease and glycaemic control are significant contributors to the overall plasma metabotype of diabetic persons, and are possible confounders in metabonomic analysis.

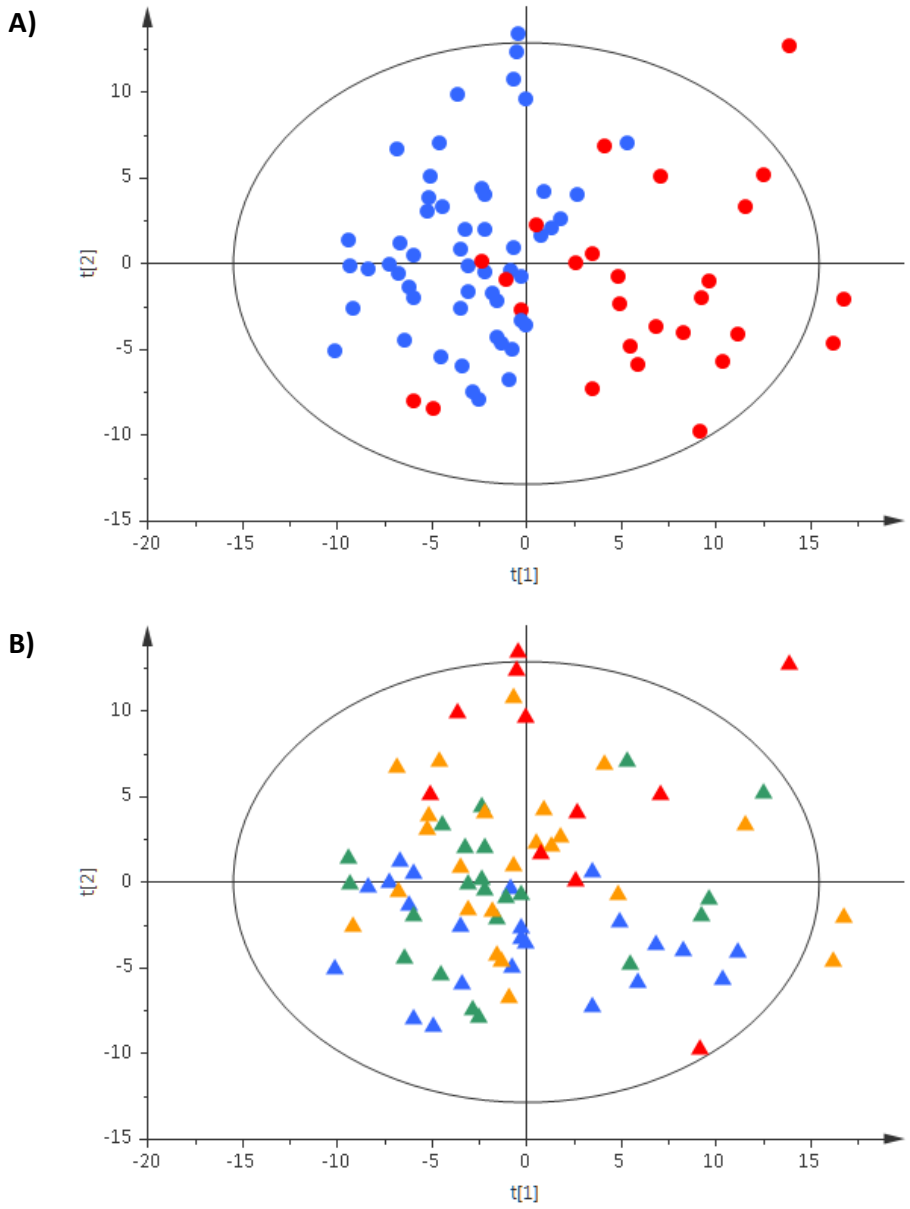


Figure 4.1 A) PCA scores plot with samples coloured by presence (red) and absence (blue) of chronic kidney disease. B) PCA scores plot with samples coloured by HbA1c levels: under 7% (blue), 7 to 8% (green), 8 to 10% (orange) and above 10% (red).

4.3.3. Metabolite Markers of DR

Fourteen metabolites were significantly different between the DR and non-DR group, based on global comparison of all DR and non-DR samples (Table 4.2). Subset analysis on the participants with HbA1c levels at 10% or below yielded eight metabolites which significantly differed between the DR and non-DR groups (Table 4.3). Further restriction to participants with HbA1c levels between 6 to 10% yielded twelve metabolites which were significantly different between DR and non-DR groups (Table 4.4). Most identified metabolite markers were concordant among two or more comparisons. 1,5-Anhydroglucitol, 1,5-gluconolactone, 2-deoxyribonic acid, gluconic acid, lactose/cellobiose and urea were consistently identified as metabolite markers in all three sets of comparative analysis (Figure 4.2). Analytical information of all identified metabolite markers are provided in Table 4.5.

The odds ratios of the metabolite markers in the basic model and adjusted values obtained from the inclusion of systolic BP, BMI, creatinine, glucose and HbA1c levels are shown in Table 4.6; and adjusted values obtained from the inclusion of triglycerides, diabetes duration, urine albumin-to-creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR) are shown in Table 4.7. Adjustment for HbA1c levels improved model performance for the majority of the metabolites, including that of 1,5-anhydroglucitol, 1,5-gluconolactone, 2-deoxyribonic acid, gluconic acid, lactose/cellobiose and urea. The inclusion of age, gender, diastolic BP, total cholesterol, HDL cholesterol and LDL cholesterol levels as co-variates gave similar odds ratios to the basic model and did not improve model performance (Table 8.5 and Table 8.6 in the appendices).

Table 4.2 Metabolite markers identified from global metabonomic analysis of all samples.

Metabolite	Fold-change*	*P*-value	FDR	AUC
1,5-Anhydroglucitol†	0.50	<0.001	0.030	0.74 (0.63–0.84)
1,5-Gluconolactone	1.45	0.001	0.058	0.71 (0.60–0.83)
2-Deoxyribonic acid	1.60	0.007	0.138	0.68 (0.56–0.79)
3,4-Dihydroxybutyric acid	1.26	0.019	0.226	0.65 (0.53–0.78)
Erythritol	1.25	0.019	0.226	0.63 (0.51–0.76)
Gluconic acid	1.57	0.001	0.037	0.72 (0.61–0.84)
Glutamine†	0.78	0.035	0.295	0.64 (0.52–0.76)
Hippuric acid	1.70	0.033	0.294	0.64 (0.52–0.76)
Lactose/Cellobiose	1.75	0.010	0.142	0.67 (0.55–0.79)
Maltose/Trehalose	1.74	0.003	0.088	0.70 (0.58–0.81)
Mannose	1.21	0.009	0.142	0.67 (0.55–0.79)
Myo-inositol	1.25	0.042	0.314	0.63 (0.51–0.76)
Ribose	1.42	0.016	0.217	0.66 (0.54–0.78)
Urea	1.35	0.004	0.093	0.69 (0.57–0.80)

* Values above 1 indicate increase in DR samples
† Classification in ROC was reversed for metabolites with fold-change less than 1

Table 4.3 Metabolite markers identified from subset of participants with HbA1c levels at 10% or below.

Metabolite	Fold-change*	*P*-value	FDR	AUC
1,5-Anhydroglucitol†	0.55	0.001	0.115	0.73 (0.61–0.85)
1,5-Gluconolactone	1.39	0.010	0.287	0.68 (0.55–0.81)
2-Deoxyribonic acid	1.55	0.029	0.459	0.65 (0.52–0.79)
alpha-Ketoisovalerate†	0.80	0.001	0.115	0.74 (0.63–0.86)
Gluconic acid	1.71	0.043	0.287	0.68 (0.55–0.81)
Hippuric acid	1.42	0.012	0.543	0.64 (0.51–0.77)
Lactose/Cellobiose	1.67	0.029	0.459	0.65 (0.52–0.79)
Urea	1.39	0.003	0.164	0.71 (0.58–0.83)

* Values above 1 indicate increase in DR samples
† Classification in ROC was reversed for metabolites with fold-change less than 1

Table 4.4 Metabolite markers identified from subset of participants with HbA1c levels between 6 to 10%.

Metabolite	Fold-change*	*P*-value	FDR	AUC
1,5-Anhydroglucitol†	0.55	0.002	0.110	0.73 (0.61–0.85)
1,5-Gluconolactone	1.41	0.010	0.231	0.69 (0.61–0.84)
2-Deoxyribonic acid	1.66	0.013	0.240	0.68 (0.55–0.81)
3-Methyl-2-oxovaleric acid	0.79	0.005	0.177	0.71 (0.58–0.82)
alpha-Ketoisovalerate†	0.78	<0.001	0.082	0.76 (0.65–0.88)
Erythritol	1.27	0.031	0.359	0.65 (0.51–0.86)
Erythronic acid	1.23	0.049	0.446	0.66 (0.52–0.35)
Gluconic acid	1.42	0.016	0.270	0.68 (0.54–0.85)
Inositol phosphate	1.21	0.039	0.396	0.65 (0.51–0.79)
Lactose/Cellobiose	1.79	0.014	0.252	0.68 (0.54–0.81)
Myo-inositol	1.30	0.025	0.309	0.66 (0.52–0.81)
Urea	1.45	0.002	0.110	0.73 (0.61–0.82)

* Values above 1 indicate increase in DR samples
† Classification in ROC was reversed for metabolites with fold-change less than 1

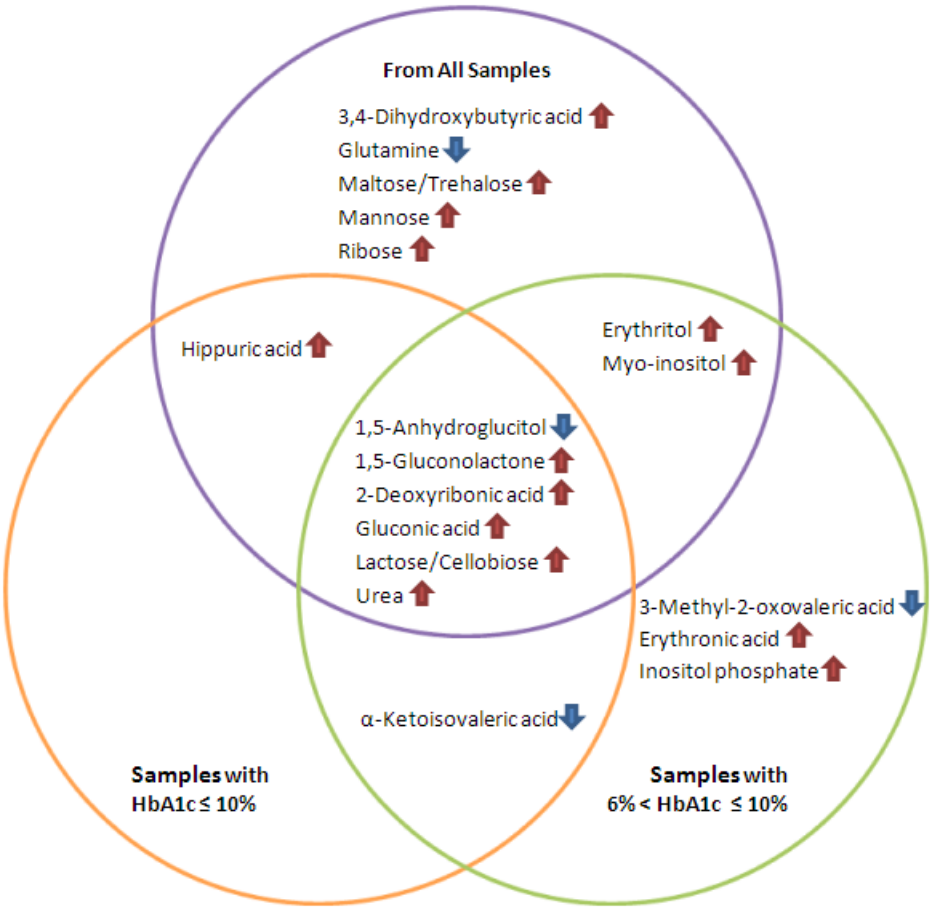


Figure 4.2 Concordance of DR metabolite markers identified from (i) all samples, (ii) subset of samples with HbA1c levels at 10% or below and (iii) subset of samples with HbA1c levels between 6 to 10%. Arrows indicate the direction of change in DR samples, compared with non-DR samples.

Table 4.5 Analytical information of DR metabolite markers.

Metabolite	Retention Index	Quantitation *m/z*	CV in QC samples
1,5-Anhydroglucitol	1882	217	5
1,5-Gluconolactone*	1910	319	15
2-Deoxyribonic acid	1689	233	13
3,4-Dihydroxybutyric acid	1447	233	8
3-Methyl-2-oxovaleric acid	1193	200	13
α-Ketoisovalerate	1129	202	15
Erythritol*	1546	217	5
Erythronic acid	1575	292	7
Gluconic acid*	2064	333	13
Glutamine	1629	246	14
Hippuric acid*	1796	105	10
Inositol phosphate	2486	318	22
Lactose/Cellobiose†	2800	361	11
Maltose/Trehalose†	2845	361	4
Mannose*	1941	319	7
Myo-inositol*	2156	305	4
Ribose*	1718	217	11
Urea	1236	189	5

* Verified with analytical standard
† Compounds unresolvable by one-dimensional GC

Table 4.6 Odds ratio of metabolite markers adjusted for systolic BP, BMI, creatinine, glucose and HbA1c.

Metabolite	Odds ratio, per s.d. increment in metabolite peak area (95% CI)					
	Basic Model	+ Systolic BP	+ BMI	+ Creatinine	+ Glucose	+ HbA1c
1,5-Anhydroglucitol	0.37 (0.19–0.65)*	0.36 (0.19–0.64)*	0.35 (0.18–0.62)*	0.38 (0.19–0.67)*	0.40 (0.21–0.71)*	0.51 (0.25–0.97)*‡
1,5-Gluconolactone	2.40 (1.38–4.65)*	2.58 (1.46–5.05)*	2.52 (1.44–4.91)*	2.63 (1.36–5.69)*	2.21 (1.25–4.35)*	2.28 (1.24–4.78)*‡
2-Deoxyribonic acid	2.08 (1.25–3.76)*	2.22 (1.32–4.11)*	2.19 (1.31–4.03)*	2.92 (1.31–7.58)*	1.97 (1.16–3.61)*§	2.14 (1.22–4.14)*‡
3,4-Dihydroxybutyric acid	1.61 (1.01–2.70)†	1.67 (1.05–2.80)*	1.63 (1.02–2.73)*	1.49 (0.88–2.61)	1.56 (0.96–2.62)†‡	1.51 (0.91–2.59)‡
3-Methyl-2-oxovaleric acid	0.81 (0.50–1.27)	0.78 (0.48–1.23)	0.75 (0.46–1.18)	0.94 (0.55–1.58)	0.68 (0.40–1.12)‡	0.44 (0.22–0.82)*‡
α-Ketoisovalerate	0.58 (0.34–0.93)*	0.55 (0.32–0.89)*	0.54 (0.31–0.88)*	0.61 (0.33–1.07)†	0.54 (0.30–0.89)*‡	0.31 (0.14–0.59)*‡
Erythritol	1.68 (1.04–2.85)*	1.75 (1.08–2.97)*	1.79 (1.10–3.07)*	1.79 (0.86–3.99)	1.64 (1.00–2.82)†‡	1.71 (1.02–3.05)†‡
Erythronic acid	1.57 (0.97–2.71)†	1.67 (1.02–2.93)†	1.59 (0.98–2.75)†	1.53 (0.71–3.46)	1.62 (0.98–2.92)†‡	1.97 (1.11–3.88)*‡
Gluconic acid	2.65 (1.50–5.33)*	2.76 (1.55–5.57)*	2.75 (1.54–5.54)*	2.83 (1.47–6.25)*	2.49 (1.30–5.38)*	2.04 (1.10–4.20)*‡
Glutamine	0.51 (0.28–0.86)*	0.52 (0.29–0.86)*	0.56 (0.31–0.92)*	0.53 (0.29–0.87)*	0.48 (0.25–0.81)*‡	0.53 (0.28–0.90)*‡
Hippuric acid	1.71 (1.05–2.98)*	1.75 (1.07–3.06)*	1.66 (1.00–2.93)†	1.61 (0.91–2.98)	1.78 (1.06–3.20)*‡	2.05 (1.16–3.97)*‡
Inositol phosphate	1.26 (0.80–2.03)	1.34 (0.85–2.15)	1.46 (0.92–2.40)§	1.30 (0.82–2.12)	1.52 (0.93–2.58)‡	1.69 (1.01–2.99)†‡
Lactose/Cellobiose	1.78 (1.10–3.01)*	1.93 (1.19–3.29)*	1.79 (1.11–3.05)*	1.68 (0.98–3.02)†	1.72 (1.05–2.96)*§	1.92 (1.13–3.51)*‡
Maltose/Trehalose	2.14 (1.29–3.84)*	2.01 (1.21–3.57)*	2.04 (1.24–3.59)*	2.09 (1.26–3.75)*	1.94 (1.07–3.74)*	1.46 (0.80–2.82)‡
Mannose	1.81 (1.09–3.27)*	1.65 (1.00–2.93)†	1.73 (1.06–3.04)*	1.88 (1.13–3.40)*§	1.62 (0.97–2.79)†§	1.34 (0.74–2.35)‡
Myo-inositol	1.60 (0.99–2.76)†	1.69 (1.04–2.95)*	1.83 (1.11–3.24)*§	1.64 (0.74–3.82)	1.56 (0.95–2.72)†‡	1.65 (0.97–3.05)†‡
Ribose	1.65 (1.02–2.80)*	1.68 (1.04–2.85)*	1.62 (1.01–2.76)†	1.52 (0.90–2.71)	1.57 (0.96–2.72)†§	1.49 (0.89–2.64)‡
Urea	1.95 (1.18–3.48)*	2.17 (1.30–3.95)*§	2.08 (1.25–3.74)*	2.12 (1.09–4.36)*	2.07 (1.20–3.91)*‡	2.50 (1.39–4.98)*‡
* P(model)< 0.05	‡ P(likelihood-ratio)<0.05					
† P(model)<0.1	§ P(likelihood-ratio)<0.1					

Table 4.7 Odds ratio of metabolite markers adjusted for triglycerides, diabetes duration, urine ACR and eGFR.

Metabolite	Odds ratio, per s.d. increment in metabolite peak area (95% CI)				
	Basic Model	+ Triglycerides	+ Diabetes Duration	+ Urine ACR	+ eGFR
1,5-Anhydroglucitol	0.37 (0.19–0.65)*	0.34 (0.16–0.61)*	0.37 (0.19–0.64)*	0.39 (0.20–0.70)*	0.37 (0.19–0.65)*
1,5-Gluconolactone	2.40 (1.38–4.65)*	2.49 (1.42–4.83)*	2.38 (1.36–4.65)*	2.24 (1.25–4.41)*	2.82 (1.43–6.31)*
2-Deoxyribonic acid	2.08 (1.25–3.76)*	2.14 (1.27–3.89)*	2.10 (1.24–3.86)*	1.91 (1.11–3.55)*	3.43 (1.49–9.20)*
3,4-Dihydroxybutyric acid	1.61 (1.01–2.70)†	1.63 (1.02–2.71)*	1.59 (0.99–2.67)†	1.49 (0.92–2.51)	1.55 (0.90–2.82)
3-Methyl-2-oxovaleric acid	0.81 (0.50–1.27)	0.75 (0.46–1.17)	0.82 (0.51–1.30)	0.89 (0.54–1.45)§	0.91 (0.53–1.55)
α-Ketoisovalerate	0.58 (0.34–0.93)*	0.55 (0.32–0.89)*	0.58 (0.34–0.94)*	0.63 (0.36–1.03)†	0.58 (0.31–1.03)†
Erythritol	1.68 (1.04–2.85)*	1.73 (1.07–2.93)*	1.67 (1.02–2.87)*	1.50 (0.89–2.62)	2.15 (0.97–5.33)†
Erythronic acid	1.57 (0.97–2.71)†	1.62 (1.01–2.80)†	1.56 (0.95–2.74)†	1.39 (0.82–2.48)	1.68 (0.80–3.88)
Gluconic acid	2.65 (1.50–5.33)*	2.75 (1.53–5.56)*	2.64 (1.49–5.31)*	2.47 (1.37–5.03)*	2.87 (1.50–6.30)*
Glutamine	0.51 (0.28–0.86)*	0.52 (0.29–0.86)*	0.51 (0.28–0.85)*	0.52 (0.28–0.88)*§	0.53 (0.29–0.88)*
Hippuric acid	1.71 (1.05–2.98)*	1.75 (1.08–3.03)*	1.69 (1.03–2.98)†	1.70 (0.98–3.09)†	1.66 (0.95–3.03)†
Inositol phosphate	1.26 (0.80–2.03)	1.41 (0.90–2.29)	1.25 (0.79–2.02)	1.27 (0.80–2.06)§	1.28 (0.81–2.07)
Lactose/Cellobiose	1.78 (1.10–3.01)*	1.84 (1.15–3.11)*	1.77 (1.09–3.02)*	1.64 (1.00–2.81)†	1.77 (1.02–3.25)†
Maltose/Trehalose	2.14 (1.29–3.84)*	2.03 (1.23–3.60)*	2.17 (1.30–3.91)*	2.05 (1.23–3.67)*	2.15 (1.29–3.87)*
Mannose	1.81 (1.09–3.27)*	1.64 (1.01–2.89)†	1.86 (1.11–3.38)*	1.81 (1.08–3.31)*§	2.00 (1.19–3.70)*§
Myo-inositol	1.60 (0.99–2.76)†	1.68 (1.03–2.91)*	1.58 (0.98–2.74)†	1.41 (0.82–2.53)	1.86 (0.84–4.58)
Ribose	1.65 (1.02–2.80)*	1.70 (1.06–2.91)*	1.63 (1.01–2.78)†	1.57 (0.96–2.70)†	1.57 (0.94–2.76)†
Urea	1.95 (1.18–3.48)*	2.07 (1.25–3.72)*	1.94 (1.17–3.47)*	1.83 (1.10–3.26)*	2.78 (1.28–6.69)*

* P(model)< 0.05 ‡ P(likelihood-ratio)<0.05
† P(model)<0.1 § P(likelihood-ratio)<0.1

4.3.4. Metabolite Markers Complementary to Known Risk Factors

To improve existing risk stratification of DR-free diabetic patients and those suffering from early stages of DR, it is relevant to identify metabolite markers of DR that are complementary to known risk factors such as poor glycaemic control and hypertension (136, 137). From Figure 4.2, we observe that hippuric acid remains a metabolite marker among the subset of participants with HbA1c levels at 10% and below. Erythritol and myo-inositol are also metabolite markers among the subset of participants with HbA1c levels between 6 to 10%. More importantly, we have identified a robust panel of metabolite markers which are not only independent of HbA1c, but also remain statistically significant when adjusted for known metabolic risk factors for DR and potentially confounding kidney disease. From Table 4.6 and Table 4.7, we observe that 1,5-gluconolactone, 2-deoxyribonic acid, gluconic acid and urea retain significant odds ratios when adjusted for systolic BP, BMI, triglycerides, blood glucose, HbA1c, duration of diabetes, urine ACR and eGFR.

4.3.5. Pathways Associated with DR

Pathway mapping on MetaboAnalyst (153) demonstrated that the pentose phosphate and galactose metabolism pathways are significantly enriched among the metabolite markers in these results. The pentose phosphate pathway was identified with 1,5-gluconolactone, gluconic acid and ribose; while galactose metabolism was identified with lactose, mannose and myo-inositol (Table 4.8). These results suggest the implications of selected metabolite classes (organic acids, polyols and sugars, oxidative stress markers and co-metabolites of microbial metabolism) in DR pathogenesis and further studies are warranted.

Table 4.8 Metabolic pathways dysregulated in DR.

Pathway	Metabolites Involved	*P*-value from enrichment analysis	FDR
Pentose phosphate pathway	1,5-Gluconolactone, Gluconic acid, Ribose	6×10^{-4}	0.044
Galactose metabolism	Lactose, Mannose, Myo-inositol	1×10^{-3}	0.092

3,4-Dihydroxybutyric acid was identified as a metabolite marker in the global comparison of all selected participants (Table 4.2) while 2-deoxyribonic was identified as a metabolite marker in both the global comparison and subset analyses (Table 4.2, Table 4.3, Table 4.4, Figure 4.2). These two metabolites have not been reported to be associated with diabetes or DR, nor are they curated in the major metabolic major databases such as KEGG (100), Reactome (154) and MetaCore. 3,4-Dihydroxybutyric acid is a urinary marker for succinate semialdehyde dehydrogenase deficiency (155) while the presence of 2-deoxyribonic acid within DNA is an indication of oxidative damage to the sugar moiety (156). Further investigations of the function of 2-deoxyribonic acid and correlation studies between free metabolite 2-deoxyribonic acid and DNA-bound 2-deoxyribonic acid are required to establish its relevance in DR pathogenesis.

Increased polyol pathway flux is one of the main driving hypotheses responsible for diabetic vascular complications (157). Increased aldose reductase activity under hyperglycemic conditions consumes NADPH. This affects regeneration of reduced glutathione, thereby induces or worsens underlying oxidative stress. While Barba *et al*. established indirect evidence of increased polyol pathway flux in the analysis of vitreous humour of DR patients (21), increases in plasma levels of sorbitol and fructose, products of the polyol pathway, were not observed among participants

with DR in our study. This is likely attributable to localisation of aldose reductase activity on glucose in susceptible cell types—in this case, retina capillary endothelium (157). As a result, this perturbation remains undetected in systemic circulation such as plasma. The pentose phosphate pathway is the main source of cellular NADPH. Liu *et al*. had identified activation of the pentose phosphate pathway in adiponectin knock-out mice on a high-fat diet (158). This was associated with disturbances in glucose homeostasis and lipid metabolism. The activation pentose phosphate pathway in our study (Table 4.8), particularly the increases in metabolites in the oxidative stage of the pentose pathway allude to involvement of oxidative stress in the pathogenesis of DR (Figure 4.3).

Erythritol is recommended as a diabetic-safe sweetener as it is a metabolically inert anti-oxidant (159). Earlier metabonomic studies have identified erythritol as a marker of diabetes (64) and impaired fasting glucose (59). In this study, plasma erythritol levels were increased in DR participants in the global comparison and among the subset of participants with HbA1c between 6 to 10% (Table 4.2, Table 4.4, Figure 4.2). In the absence of dietary survey data, it is unclear if the observed trends in erythritol levels were due to increased consumption of sweeteners or DR-related metabolic dysregulation. Mannose, a metabolite in galactose metabolism was previously identified as a marker for impaired fasting glucose (58, 59) and diabetes (58, 64). In this study, plasma mannose levels were increased in DR participants in the global comparison (Table 4.2, Figure 4.2). Therefore future studies targeting the quantitation of polyols and sugars are pertinent as metabolites in the pentose phosphate and galactose pathways are potential clinical metabolite markers of DR. Dietary information should also be incorporated

to allow the distinct identification of metabolite markers associated with DR from metabolic perturbations that directly result from nutritional factors.

There is accumulating evidence to support the contribution of altered gut microbiota to obesity and diabetes (160). More recently, metagenome-wide association studies demonstrated the association of gut microbial dysbiosis with type 2 diabetes (161, 162). Hippuric acid is a co-metabolite of gut microbiota derived from benzoic acid (163). The identification of hippuric acid as a metabolite marker (Table 4.2, Table 4.3) suggests that altered gut microbiota could contribute to DR pathogenesis. Further studies would be required to determine the roles of gut microbiota composition and microbial co-metabolism in DR and other vascular complications of diabetes.

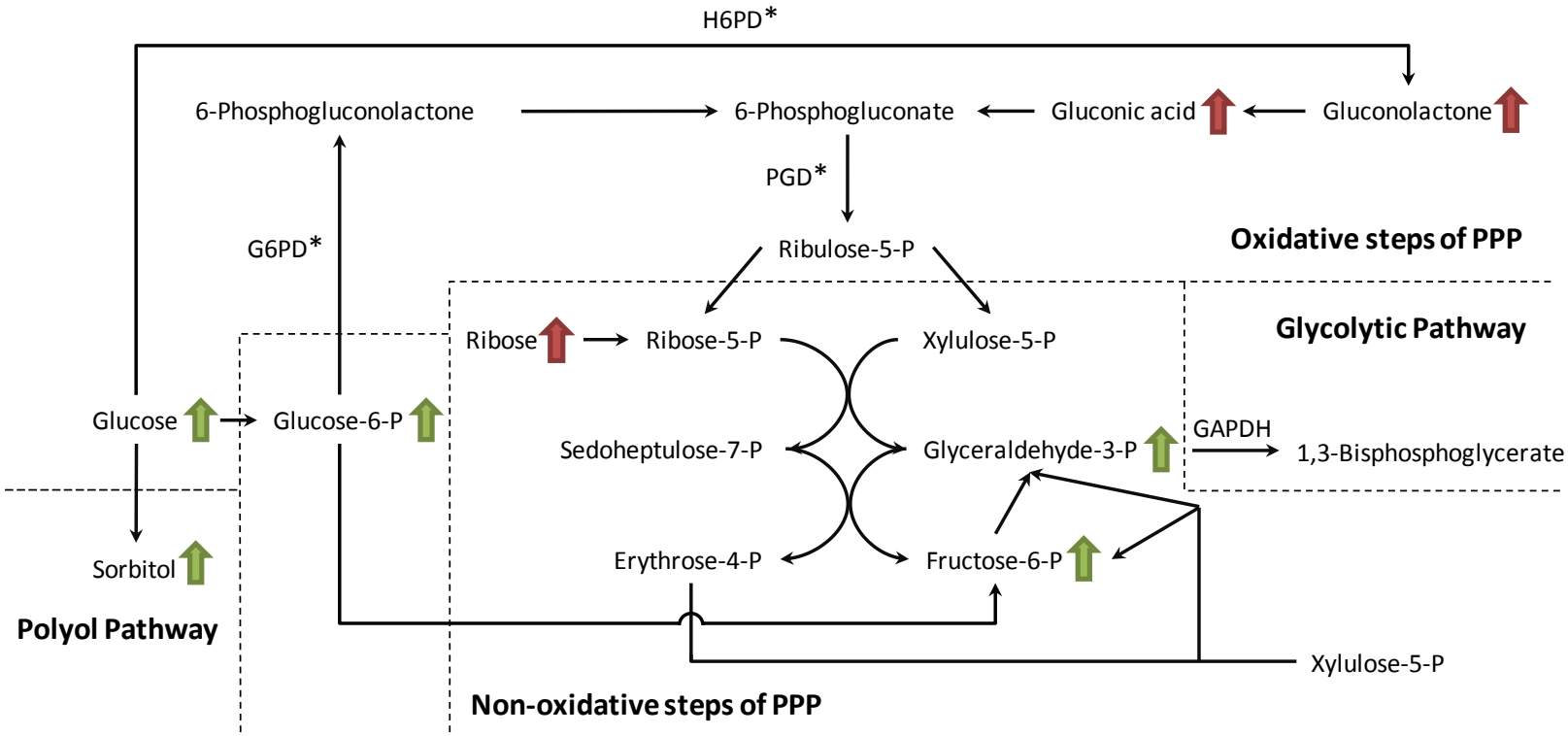


Figure 4.3 Activation of the pentose phosphate pathway identified from elevated 1,5-gluconolactone, gluconic acid and ribose (red arrows) levels in DR patients. Reactions marked with an asterisk produce NADPH. Green arrows denote the accumulation of glycolytic metabolites leading activation of the polyol, hexosamine, protein kinase C and AGE pathways in the unifying mechanism responsible for diabetic complications, as proposed by M. Brownlee.

4.3.6. Non-Replicable Findings

There are however, several metabolite trends which have been associated with type 2 diabetes or DR that were not replicated in this study.

Increases in plasma/sera branched chain (62-64) and aromatic (60-63) amino acids levels were reported as markers and predictors of type 2 diabetes, while glycine levels were found to be negatively correlated with diabetes (63, 65). Unlike earlier findings, the levels of branched chain amino acids and glycine were not significantly different between DR and non-DR groups in this study (Table 4.9). This contrast is likely due to the definition and choice of the control group. In the cited studies, the control samples were derived from healthy persons, while this study's controls were sampled from diabetic persons who could have pre-existing increased catabolism of branched chain and aromatic amino acids. In addition, the majority of participants in this study had been prescribed with diabetes-controlling medication (Table 4.1). These medications might have helped to lower the magnitude of perturbations in amino acid metabolism compared with untreated diabetes.

3-Hydroxybutyric acid, a ketone body, is a marker for impaired glucose tolerance (57) and type 2 diabetes (64, 65). Although Li *et al.* reported 3-hydroxybutyric acid as marker of DR (22), this study did not detect an increase in plasma levels of this ketone body in DR participants (Table 4.9). 2-Hydroxybutyric acid, a derivative of α-ketobutyrate, was previously reported as a marker (58, 65) and predictor (164) for diabetes, but its levels were not significantly different between DR and non-DR groups in this study (Table 4.9). These non-replicable trends imply that the plasma metabotype of DR is not a mere extension or subset of the metabotype of

diabetes. Li *et al.* also reported lower of plasma linoleic and arachidonic acids in DR participants. While this is in line with the view that arachidonic acid generally produces pro-inflammatory eicosanoids (165), we did not find significant differences in the levels of linoleic acid and arachidonic acid between DR and non-DR participants in this study (Table 4.9). The differences between this study's findings and those reported in previous literature could be due to unaccounted confounders. For example, Li *et al.* did not consider BP and diabetes duration. The increase in plasma aspartic acid levels identified by Li *et al.* (22) might potentially be a marker of co-morbid diabetic nephropathy (166).

Omega-3 polyunsaturated fatty acids (PUFAs) are thought to aid in diabetes prevention but meta-analyses on omega-3 fatty acids intake show heterogeneous results, stratified by study locale (167). The omega-3 PUFAs, docosahexaenoic and eicosapentaenoic acid were detected in this study but there were no significant differences in the plasma levels of these PUFAs between DR and non-DR groups (Table 4.9). Similarly, Mäkinen *et al.* found that the baseline serum docosahexaenoic could not predict progressive kidney disease among patients with type 1 diabetes (168). Although omega-3 PUFAs have been shown to protect against DR in rodent models (169-171), their efficacy in preventing vascular complications of diabetes can only be proven through randomised controlled trials.

Table 4.9 Diabetes or DR-associated metabolites in literature presenting no significant changes between DR and non-DR participants in this study.

Compound Class	Metabolite	Fold-Change	*P*-value	Retention Index	CV in QC samples
Amino Acid	Aspartic acid	0.82	0.13	1530	15
	Glycine	0.99	0.94	1317	9
	Isoleucine	0.96	0.67	1170	14
	Leucine	0.94	0.74	1150	13
	Phenylalanine	0.94	0.26	1624	11
	Tryptophan	0.97	0.33	2178	9
	Tyrosine	0.85	0.14	1871	26
	Valine	0.92	0.19	1083	11
Hydroxy Acid	2-Hydroxybutyric acid	0.98	0.80	1136	8
	3-Hydroxybutyric acid	1.27	0.39	1166	8
PUFA	Arachidonic acid	1.05	0.38	2353	18
	Docosahexaenoic acid	0.98	0.77	2538	22
	Eicosapentaenoic acid	0.92	0.20	2376	24
	Linoleic acid	0.98	0.73	2200	11
	Oleic acid	0.98	0.87	2208	4

4.3.7. Confounding Factors in Pathway Analysis.

GC-MS is widely used in metabonomic analysis as EI ionisation is highly robust and reproducible, enabling the use of spectral libraries for metabolite identification. However, this platform is prone to confounding factors introduced through the sample derivatisation process or thermal degradation of metabolites at elevated temperatures. For example, arginine is detected as carbodiimide (172) (Kovat's retention index of 972 in this study) and cyclic lactones are known to partially decompose to their respective sugar acids when analysed by GC-MS (173). We are however, confident of the identification of 1,5-gluconolactone as a marker of DR as it demonstrated good reproducibility in QC samples (15% CV, Table 4.5) and yielded a linear peak area response with concentration in standards (Pearson's r = 0.99, 1 to 25 µg/mL). Fatty acids detected by GC-MS reflect the total sum present in triglycerides, diacyglycerols and other lipids (6), and are not an accurate reflection of circulating free fatty acid concentrations. Future investigations on lipid metabolism in DR would be better investigated using a targeted lipidomics approach.

4.3.8. Study Limitations

This study focused on the Indian (South Asian) ethnic group as the prevalence of diabetes and age-adjusted DR in Indians is the highest among the major ethnic groups of Singapore (174). As this study design only captures a snapshot of plasma metabotypes, some of the identified markers are potentially a representation of short-term metabolic perturbations instead of chronic risk factors associated with development of DR. For example, 1,5-anhydroglucitol, a marker for short-term glycaemic control (lowered values indicate hyperglycaemia and vice-versa) was found to be lowered in DR participants in all comparative analysis.

4.4. Summary for Metabotyping of Diabetic Retinopathy

Using global metabonomic profiling, this study identified plasma metabolite signatures of DR which are resolvable from pre-existing diabetes and remain robust in the presence of confounders. The metabolite markers 1,5-gluconolactone, urea, gluconic acid and 2-deoxyribonic acid were significantly different between DR and non-DR groups, even after adjustments for clinicopathological variables. The pentose phosphate pathway was identified as a key metabolic dysregulation associated with DR and points to the involvement of oxidative stress towards disease pathogenesis. The identification of hippuric acid, a microbial co-metabolite suggests that the dysregulation of the microbiome could be another risk factor for DR.

5. Investigation of SWATH™ MS for Metabonomic Analysis

5.1. Introduction to SWATH™ Technique

SWATH™ MS is a new data acquisition mode recently introduced on Q-TOF instruments (69). Unlike in IDA where only a limited number of precursor *m/z* of high signal intensities are selected for fragmentation, SWATH™ MS acquires continuous MS/MS data in each duty cycle. This is achieved by modifying the radio frequency and direct current on the quadrupole to select isolation windows that are 10 to 25 Da wide. These selection windows are applied in steps across the entire precursor mass range within one duty cycle. Thus SWATH™ acquisition captures MS/MS information from precursors of low signal intensity that would have been missed by IDA. Quantitative information in SWATH™-acquired data are obtained via high resolution XICs of MS/MS product ions recorded in the respective isolation windows where the parent *m/z* of the compounds of interest lie. Unlike conventional multiple-reaction-monitoring (MRM) quantitation methods, SWATH™ MS captures quantitative data of all detectable species in a sample in a permanent archive. These data files could be mined to validate the findings of discovery -omic studies without the need to develop MRM parameters for each compound, or to identify new candidates in the same metabolic pathways as newly-discovered biomarkers.

Due to the width of the isolation window, quantitation by SWATH™ MS is susceptible to crosstalk from co-eluting precursor ions that produce fragments

with a similar *m/z* (even at high resolution). In proteomics, this issue can be circumvented by (i) the application of FDR filters on MS/MS spectra, (ii) the selection of unique ion signatures to proteins-of-interest (175), and (iii) the selection of other multiply-charged *m/z* species of the same sequence for quantitation.

To the best of our knowledge, there is no literature available evaluating the use of SWATH™ MS for the analysis of underivatised endogenous metabolites. This chapter investigated the performance of SWATH™ MS in metabolite quantitation alongside conventional MRM. The metabolites evaluated are betaine, carnitine, glycerophoshocholine, phosphocholine, taurine, sarcosine and UDP-*N*-acetylglucosamine. These metabolites were either reported to be osmoprotectants from literature (120, 121, 176) and/or found to be dysregulated in this study on IOBA-NHC cells under hyperosmotic stress earlier in this thesis.

5.2. MRM and SWATH™ Quantitation Methods

Metabolite standards were purchased from Sigma Aldrich. Standard solutions were prepared in the range of 1 to 200 ppb (pg/µL) in 4:1 ACN/water. Two MRM transitions were optimised for betaine, carnitine, glycerophoshocholine, phosphocholine, taurine and sarcosine via direct infusion. The transitions were optimised for DP, CE and collision exit potential (CXP). Metabolite standards were then analysed by MRM and SWATH™ MS.

Triplicate injections were performed for each concentration. Separation in both quantitation methods was performed on an Agilent 1290 Infinity chromatography

system (Agilent Technologies) using an Atlantis HILIC silica 2.1 × 100 mm 3 µm column (Waters Corporation). The mobile phase was 5 mM ammonium formate (eluent A) and 5 mM ammonium formate (eluent B). The linear gradient profile was 100% B from 0 to 1 min, 20% B from 12 to 14 min, and 100% B from 15 to 19 min. The flow rate was 0.4 mL/min and the injection volume was 10 µL for all standards.

MRM acquisition was performed on a QTrap 6500 ion trap mass spectrometer in the positive ionisation polarity using optimised DP, CE and CXP parameters developed for each compound. The dwell time of each transition was 25 ms. For each metabolite, the transition with the greater intensity was used to plot a calibration curve in MultiQuant (AB Sciex).

SWATH™ acquisition was performed with a TripleTOF 5600 Q-TOF mass spectrometer. Betaine, carnitine, glycerophoshocholine, phosphocholine, taurine and sarcosine standards were analysed in the positive ionisation polarity, while UDP-*N*-acetylglucosamine standards were analysed in the negative ionisation polarity. The Q1 mass range was scanned from *m/z* 50 to 800 in steps of 25 Da. MS/MS data in each step was acquired from *m/z* 25 to 800. The accumulation time for each step was 50 ms and the total duty cycle time was 700 ms. The source voltage was 5 kV and declustering potential was 80 V. CE was ramped from 20 to 40 V. MS/MS spectra of the metabolite standards were also acquired to determine the accurate mass of their product ions. The accurate mass of the product ions corresponding to the MRM transition of greater intensity were extracted from their respective isolation windows with MultiQuant (AB Sciex).

5.3. Results and Discussion: Comparison of SWATH™ with MRM

5.3.1. LOQ and Linearity of SWATH™ Quantitation

The optimised MRM parameters of the metabolite panel are shown in Table 5.1. The isolation windows and MS/MS extraction masses of these metabolites are shown in Table 5.2. Pearson's *r* values of the calibration curve and the lower limit of quantitation (LOQ) obtained from MRM and SWATH™ quantitation of the metabolite panel are shown in Table 5.3. The LOQ in this study was defined as the lowest concentration at which the metabolite was detectable with a CV of 20% or less among triplicates, and with an error of 20% or less between actual and calculated concentrations. The LOQ was expressed as nanograms of metabolites injected on column. Compared with MRM, SWATH™ quantitation yielded poorer (higher) LOQ for all metabolites. The *r* values of the calibration curve for taurine was worse (lower) in SWATH™ quantitation.

Table 5.1 Optimised MRM parameters of the metabolite panel.

Compound	RT (min)	Q1 (m/z)	Q3 (m/z)	DP (V)	CE (V)	CXP (V)
Betaine	5.06	118.1	58.0	90	40	10
Carnitine	6.48	162.1	103.0	80	23	12
Glycerophoshocholine	6.60	258.1	104.2	70	21	5
Sarcosine	4.96	90.0	44.2	35	18	7
Taurine	3.26	126.0	108.0	60	16	10

Table 5.2 Isolation windows and MS/MS extraction masses of selected metabolites

Metabolite	Isolation window (m/z range)	Extraction Mass (m/z)
Betaine	99.0–125.0	58.068
Carnitine	149.0–175.0	103.039
Glycerophoshocholine	249.0–275.0	104.107
Sarcosine	74.0–100.0	44.053
Taurine	124.0–150.0	108.010

Table 5.3 Comparison of LOQ and Pearson's *r* obtained from MRM and SWATH™ quantitation of the metabolite panel.

Metabolite	LOQ (ng)		Pearson's *r*	
	MRM	SWATH™	MRM	SWATH™
Betaine	0.01	0.20	0.83	0.84
Carnitine	0.01	0.10	1.00	0.97
Glycerophoshocholine	0.01	0.10	0.98	0.92
Sarcosine	0.05	0.20	0.99	0.96
Taurine	0.10	0.50	0.98	0.87

5.3.2. Correlation of Extracted Ions

The top five fragments with the greatest intensity for UDP-*N*-acetylglucosamine are *m/z* 272.961, 282.037, 362.014, 384.983 and 403.001. Their peak areas over the concentration range of 10 to 100 ppb are shown in Figure 5.1. The peak area of *m/z* 403.001 in the investigated concentration range does not correlate with that of other fragments. This shows that SWATH™ quantitation is susceptible to interference in the background due the large *m/z* window applied in ion selection by the quadrupole. Compared with SWATH™ proteomics, the issue is aggravated in the analysis of low molecular weight metabolites. To perform accurate quantitation using SWATH™ MS in metabonomic analysis, two or more accurate mass fragments of a compound should be selected and their peak areas checked for linear correlation in standard injections. Interference could be reduced by decreasing the width of the isolation window, but at the expense of duty cycle time or sensitivity.

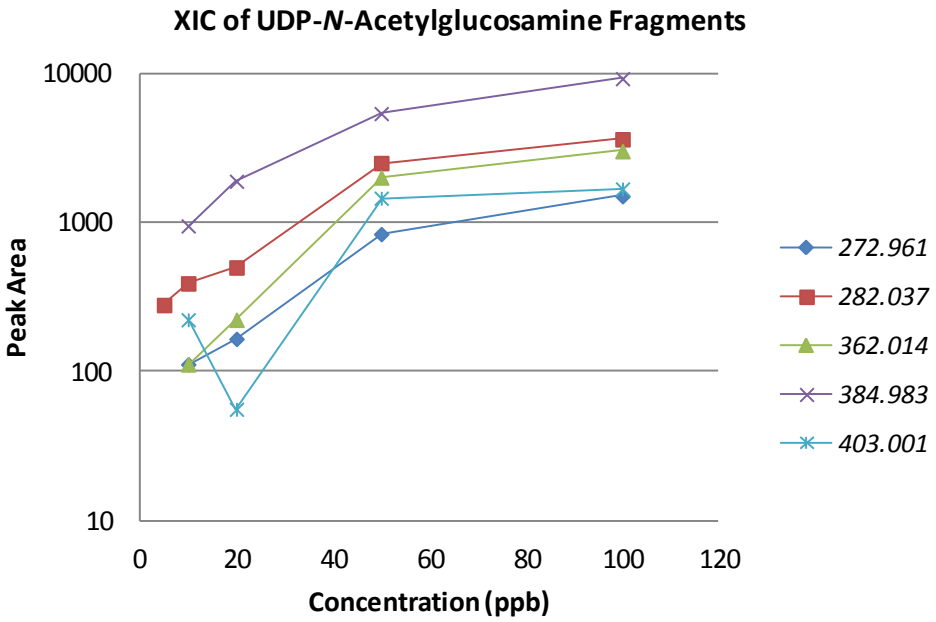


Figure 5.1 Peak areas of UDP-*N*-acetylglucosamine fragments in standard injections from 1 to 100 ppb.

5.4. Feasibility of SWATH™ MS in Metabonomic Analysis

Although the SWATH™ workflow has been applied in global proteomic analysis with free-access or in-house acquired libraries (73, 177), there are limitations to implementing SWATH™ acquisition in global metabonomic analysis. Firstly, there is no commercially-available software that is capable of (i) automated extraction of a large number of isolation window-fragment ion pairs from raw data or (ii) construction of sample-specific isolation window-fragment ion libraries from IDA acquisitions. Given the current limitation in informatics, SWATH™ MS is best limited for quantitation of known metabolites with validated MS/MS extraction masses. As SWATH™ MS has a higher LOQ than MRM, it may not be suitable for analysing small volume samples such as tear fluid.

6. Conclusion and Future Outlook

This thesis reports the first comprehensive characterisation of the human tear metabolome using IDA LC-MS/MS with RP and HILIC chromatography coupled to the positive and negative ionisation polarities. A total of 60 metabolites were identified, and 44 were detected for the first time in tear fluid. These results demonstrate the applicability of tear fluid metabolites in studying eye diseases. Several challenges were faced in metabolite identification; these apply to any experiment in global metabonomic profiling on an LC-MS/MS platform. Some examples of possible misannotations that are described in this thesis, such as in-source fragmentation of inosine, are also mentioned in a recent publication that examined such misannotations in detail (178).

Besides the aforementioned ionisation artefacts, non-biological artefacts derived from sample collection and extraction processes, such as coatings of blood collection tubes, surfactants and detergents may also contribute to *m/z*-RT features in metabonomic analysis (179). Unlike in proteomic analysis where data quality can be assessed by the number of identified proteins, the number of features obtained in LC-MS/MS-based metabonomics does not reflect the coverage of the metabolome. Mahieu *et al*. have recently developed a workflow for credentialing *m/z*-RT features using *E. coli* cultures grown in regular (control) and ^{13}C-enriched media (180). Features of biological origin were identified from isotope-intensity analysis of standing samples consisting of control: ^{13}C-labeled mixed at specific ratios. The use of such credentialing workflows would provide a common benchmark in assessing metabolome coverage in LC-MS/MS-based metabonomics.

The second chapter of this thesis identified dose-dependent accumulation of glycerophosphcholine with increasing hyperosmotic stress using an *in vitro* model of dry eye disease. With the integration of proteomic data, *O*-linked β-*N*-acetylglucosamine glycosylation was identified as a potential therapeutic target for dry eye disease. Experiments targeting the activity of choline transporters and enzymes in the synthesis pathway of glycerophosphcholine could also be performed to establish their effects on intracellular glycerophosphocholine levels. The levels *O*-linked β-*N*-acetylglucosamine glycosylation could also be modulated to investigate its effects on cell survival. It is also pertinent to investigate the tear fluid concentration of glycerophosphocholine, UDP-*N*-acetylglucosamine and other perturbed metabolites in healthy volunteers and patients with dry eye to establish their relevance as biomarkers of dry eye or ocular surface inflammation.

Lastly, plasma metabotyping of DR using GC-MS found metabolite markers of DR that were independent of known risk factors. These findings should be validated in a larger cohort involving participants from other ethnic groups. Collection of fasting blood samples is preferred over random sampling to reduce variability. A prospective cohort study is needed to identify patterns in metabolic perturbations that could be predictive of the rate of disease progression. In such a study, diabetic patients with no DR and healthy volunteers should be recruited together. Plasma samples collected at the start of the study should be performed to establish baseline levels of plasma metabolites. Should resources permit, the participants should be examined at multiple follow-ups (e.g. every two years up to ten years). It is important to first establish the presence of distinct clusters of disease progression that were first identified by Zavrelova *et al*. (142), before investigating the correlations between such clusters with the plasma metabonomic profiles.

7. References

1. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica. 1999;29(11):1181-9.
2. Fiehn O. Metabolomics – the link between genotypes and phenotypes. Plant Mol Biol. 2002;48(1):155-71.
3. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. Cell. 2008;134(5):714-7.
4. Lindon JC, Nicholson JK. Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery. TrAC, Trends Anal Chem. 2008;27(3):194-204.
5. Halket JM, Przyborowska A, Stein SE, Mallard WG, Down S, Chalmers RA. Deconvolution gas chromatography/mass spectrometry of urinary organic acids – potential for pattern recognition and automated identification of metabolic disorders. Rapid Commun Mass Spectrom. 1999;13(4):279-84.
6. Xu F, Zou L, Ong CN. Experiment-originated variations, and multi-peak and multi-origination phenomena in derivatization-based GC-MS metabolomics. TrAC, Trends Anal Chem. 2010;29(3):269-80.
7. Trufelli H, Palma P, Famiglini G, Cappiello A. An overview of matrix effects in liquid chromatography–mass spectrometry. Mass Spectrom Rev. 2011;30(3):491-509.
8. Liu G, Ji QC, Arnold ME. Identifying, evaluating, and controlling bioanalytical risks resulting from nonuniform matrix ion suppression/enhancement and nonlinear liquid chromatography−mass spectrometry assay response. Anal Chem. 2010;82(23):9671-7.
9. Theodoridis G, Gika HG, Wilson ID. LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. TrAC, Trends Anal Chem. 2008;27(3):251-60.
10. Dunn W, Erban A, Weber RM, Creek D, Brown M, Breitling R, et al. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics. 2013;9(1):44-66.
11. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: A review. Clin Cancer Res. 2009;15(2):431-40.
12. Kaddurah-Daouk R, Krishnan KRR. Metabolomics: A global biochemical approach to the study of central nervous system diseases. Neuropsychopharmacology. 2008;34(1):173-86.
13. Robertson DG, Watkins PB, Reily MD. Metabolomics in toxicology: Preclinical and clinical applications. Toxicol Sci. 2011;120(suppl 1):S146-S70.
14. Human Metabolome Database. Human metabolome database statistics 2015 [cited 16 June 2015]. Available from: http://www.hmdb.ca/statistics.
15. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Gautam B, et al. Hmdb: A knowledgebase for the human metabolome. Nucleic Acids Res. 2009;37(suppl 1):D603-D10.
16. Fris M, Tessem M-B, Čejková J, Midelfart A. The effect of single and repeated UVB radiation on rabbit cornea. Graef Arch Clin Exp. 2006;244(12):1680-7.
17. Young SP, Nessim M, Falciani F, Trevino V, Banerjee SP, Scott RAH, et al. Metabolomic analysis of human vitreous humor differentiates ocular inflammatory disease. Mol Vis. 2009;15:1210-7.

18. Galbis-Estrada C, Martinez-Castillo S, Morales JM, Vivar-Llopis B, Monleón D, Díaz-Llopis M, et al. Differential effects of dry eye disorders on metabolomic profile by ^{1}H nuclear magnetic resonance spectroscopy. Biomed Res Int. 2014:e542549.
19. Galbis-Estrada C, Pinazo-Durán MD, Martínez-Castillo S, Morales JM, Monleón D, Zanon-Moreno V. A metabolomic approach to dry eye disorders. The role of oral supplements with antioxidants and omega 3 fatty acids. Mol Vis. 2015;21:555-67.
20. Mayordomo-Febrer A, López-Murcia M, Morales-Tatay JM, Monleón-Salvado D, Pinazo-Durán MD. Metabolomics of the aqueous humor in the rat glaucoma model induced by a series of intracamerular sodium hyaluronate injection. Exp Eye Res. 2015;131(0):84-92.
21. Barba I, Garcia-Ramírez M, Hernández C, Alonso MA, Masmiquel L, García-Dorado D, et al. Metabolic fingerprints of proliferative diabetic retinopathy: An ^{1}H-NMR–based metabonomic approach using vitreous humor. Invest Ophthalmol Vis Sci. 2010;51(9):4416-21.
22. Li X, Luo X, Lu X, Duan J, Xu G. Metabolomics study of diabetic retinopathy using gas chromatography-mass spectrometry: A comparison of stages and subtypes diagnosed by western and chinese medicine. Mol Biosyst. 2011;7(7):2228-37.
23. Brown JCC, Sadler PJ, Spalton DJ, Juul SM, Macleod AF, Sönksen PH. Analysis of human aqueous humour by high resolution ^{1}H NMR spectroscopy. Exp Eye Res. 1986;42(4):357-62.
24. Yanshole VV, Snytnikova OA, Kiryutin AS, Yanshole LV, Sagdeev RZ, Tsentalovich YP. Metabolomics of the rat lens: A combined LC-MS and NMR study. Exp Eye Res. 2014;125(0):71-8.
25. Nakatsukasa M, Sotozono C, Shimbo K, Ono N, Miyano H, Okano A, et al. Amino acid profiles in human tear fluids analyzed by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. Am J Ophthalmol. 2011;151(5):799-808.e1.
26. Zhou L, Beuerman RW. Tear analysis in ocular surface diseases. Prog Retin Eye Res. 2012;31(6):527-50.
27. Grus FH, Podust VN, Bruns K, Lackner K, Fu S, Dalmasso EA, et al. Seldi-TOF-MS proteinchip array profiling of tears from patients with dry eye. Invest Ophthalmol Vis Sci. 2005;46(3):863-76.
28. de Souza G, de Godoy L, Mann M. Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors. Genome Biol. 2006;7(8):R72.
29. Zhou L, Beuerman RW, Chan CM, Zhao SZ, Li XR, Yang H, et al. Identification of tear fluid biomarkers in dry eye syndrome using iTRAQ quantitative proteomics. J Proteome Res. 2009;8(11):4889-905.
30. Shine WE, McCulley JP. Polar lipids in human meibomian gland secretions. Curr Eye Res. 2003;26(2):89-94.
31. Zhou L, Zhao SZ, Koh SK, Chen L, Vaz C, Tanavde V, et al. In-depth analysis of the human tear proteome. J Proteomics. 2012;75(13):3877-85.
32. Chen J, Green-Church KB, Nichols KK. Shotgun lipidomic analysis of human meibomian gland secretions with electrospray ionization tandem mass spectrometry. Invest Ophthalmol Vis Sci. 2010;51(12):6220-31.
33. Butovich IA. Lipidomic analysis of human meibum using HPLC–MS^{n}. In: Armstrong D, editor. Lipidomics. Methods in molecular biology™. 579. New York: Humana Press; 2010. p. 221-46.

34. Butovich IA. On the lipid composition of human meibum and tears: Comparative analysis of nonpolar lipids. Invest Ophthalmol Vis Sci. 2008;49(9):3779-89.
35. Choy CKM, Cho P, Chung W-Y, Benzie IFF. Water-soluble antioxidants in human tears: Effect of the collection method. Invest Ophthalmol Vis Sci. 2001;42(13):3130-4.
36. Gogia R, Richer SP, Rose RC. Tear fluid content of electrochemically active components including water soluble antioxidants. Curr Eye Res. 1998;17(3):257-63.
37. Van Haeringen NJ, Glasius E. Collection method dependant concentrations of some metabolites in human tear fluid, with special reference to glucose in hyperglycaemic conditions. Albrecht Von Graefes Arch Klin Exp Ophthalmol. 1977;202(1):1-7.
38. Farkas Á, Vámos R, Bajor T, Müllner N, Lázár Á, Hrabá A. Utilization of lacrimal urea assay in the monitoring of hemodialysis: Conditions, limitations and lacrimal arginase characterization. Exp Eye Res. 2003;76(2):183-92.
39. Baca JT, Taormina CR, Feingold E, Finegold DN, Grabowski JJ, Asher SA. Mass spectral determination of fasting tear glucose concentrations in nondiabetic volunteers. Clin Chem. 2007;53(7):1370-2.
40. Taormina CR, Baca JT, Asher SA, Grabowski JJ, Finegold DN. Analysis of tear glucose concentration with electrospray ionization mass spectrometry. J Am Soc Mass Spectrom. 2007;18(2):332-6.
41. Pescosolido N, Imperatrice B, Koverech A, Messano M. L-carnitine and short chain ester in tears from patients with dry eye. Optom Vis Sci. 2009;86(2):E132-E8.
42. Trope GE, Rumley AG. Catecholamine concentrations in tears. Exp Eye Res. 1984;39(3):247-50.
43. Struck H-G, Wicht A, Pönicke K, Lautenschläger C, Lübbe D. Histamine content in tears in patients with allergic rhinoconjunctivitis. Der Ophthalmologe. 1998;95(4):241-6.
44. Pintor J, Carracedo G, Alonso C, Bautista A, Peral A. Presence of diadenosine polyphosphates in human tears. Pflug Arch Eur J Phy. 2002;443(3):432-6.
45. Mendelsohn ME, Abramson DH, Senft S, Servodidio CA, Gamache PH. Uric acid in the aqueous humor and tears of retinoblastoma patients. J AAPOS. 1998;2(6):369-71.
46. Speek AJ, van Agtmaal EJ, Saowakontha S, Schreurs WHP, van Haeringen NJ. Fluorometric determination of retinol in human tear fluid using high-performance liquid chromatography. Curr Eye Res. 1986;5(11):841-5.
47. Lemp MA. The definition and classification of dry eye disease: Report of the definition and classification subcommittee of the international dry eye workshop (2007). Ocul Surf. 2007;5(2):75-92.
48. Li D-Q, Chen Z, Song XJ, Luo L, Pflugfelder SC. Stimulation of matrix metalloproteinases by hyperosmolarity via a JNK pathway in human corneal epithelial cells. Invest Ophthalmol Vis Sci. 2004;45(12):4302-11.
49. Li D-Q, Luo L, Chen Z, Kim H-S, Song XJ, Pflugfelder SC. JNK and erk map kinases mediate induction of il-1[beta], tnf-[alpha] and il-8 following hyperosmolar stress in human limbal epithelial cells. Exp Eye Res. 2006;82(4):588-96.
50. Luo L, Li D, Pflugfelder SC. Hyperosmolarity-induced apoptosis in human corneal epithelial cells is mediated by cytochrome c and MAPK pathways. Cornea. 2007;26(4):452-60.

51. Albertsmeyer A-C, Kakkassery V, Spurr-Michaud S, Beeks O, Gipson IK. Effect of pro-inflammatory mediators on membrane-associated mucins expressed by human ocular surface epithelial cells. Exp Eye Res. 2010;90(3):444-51.
52. Png E, Samivelu GK, Yeo SH, Chew J, Chaurasia SS, Tong L. Hyperosmolarity-mediated mitochondrial dysfunction requires transglutaminase-2 in human corneal epithelial cells. J Cell Physiol. 2011;226(3):693-9.
53. Versura P, Profazio V, Schiavi C, Campos EC. Hyperosmolar stress upregulates HLA-DR expression in human conjunctival epithelium in dry eye patients and in vitro models. Invest Ophthalmol Vis Sci. 2011.
54. Cammarata PR, Schafer G, Chen S-W, Guo Z, Reeves RE. Osmoregulatory alterations in taurine uptake by cultured human and bovine lens epithelial cells. Invest Ophthalmol Vis Sci. 2002;43(2):425-33.
55. Congdon NG, Friedman DS, Lietman T. Important causes of visual impairment in the world today. JAMA: J A Med Assoc. 2003;290(15):2057-60.
56. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. Lancet. 2010;376(9735):124-36.
57. Ho JE, Larson MG, Vasan RS, Ghorbani A, Cheng S, Rhee EP, et al. Metabolite profiles during oral glucose challenge. Diabetes. 2013;62(8):2689-98.
58. Xu F, Tavintharan S, Sum CF, Woon K, Lim SC, Ong CN. Metabolic signature shift in type 2 diabetes mellitus revealed by mass spectrometry-based metabolomics. J Clin Endocrinol Metab. 2013;98(6):E1060-E5.
59. Menni C, Fauman E, Erte I, Perry JRB, Kastenmüller G, Shin S-Y, et al. Biomarkers for type 2 diabetes and impaired fasting glucose using a non-targeted metabolomics approach. Diabetes. 2013;62(12):4270-6.
60. Würtz P, Mäkinen V-P, Soininen P, Kangas AJ, Tukiainen T, Kettunen J, et al. Metabolic signatures of insulin resistance in 7,098 young adults. Diabetes. 2012;61(6):1372-80.
61. Tai ES, Tan MLS, Stevens RD, Low YL, Muehlbauer MJ, Goh DLM, et al. Insulin resistance is associated with a metabolic profile of altered protein metabolism in chinese and asian-indian men. Diabetologia. 2010;53(4):757-67.
62. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. Nat Med. 2011;17(4):448-53.
63. Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost H-G, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes. 2013;62(2):639-48.
64. Suhre K, Meisinger C, Döring A, Altmaier E, Belcredi P, Gieger C, et al. Metabolic footprint of diabetes: A multiplatform metabolomics study in an epidemiological setting. PLoS ONE. 2010;5(11):e13953.
65. Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. PLoS ONE. 2010;5(12):e15234.
66. Li X, Xu Z, Lu X, Yang X, Yin P, Kong H, et al. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus. Anal Chim Acta. 2009;633(2):257-62.
67. Kuo JZ, Wong TY, Rotter JI. Challenges in elucidating the genetics of diabetic retinopathy. JAMA Ophthalmology. 2014;132(1):96-107.
68. Lavanya R, Jeganathan VSE, Zheng Y, Raju P, Cheung N, Tai ES, et al. Methodology of the Singapore indian chinese cohort (sicc) eye study: Quantifying

ethnic variations in the epidemiology of eye diseases in asians. Ophthalmic Epidemiol. 2009;16(6):325-36.
69. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. Mol Cell Proteomics. 2012;11(6):O111.016717.
70. Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, Wolski WE, et al. Building high-quality assay libraries for targeted analysis of SWATH MS data. Nat Protocols. 2015;10(3):426-41.
71. Haverland NA, Fox HS, Ciborowski P. Quantitative proteomics by SWATH-MS reveals altered expression of nucleic acid binding and regulatory proteins in hiv-1-infected macrophages. J Proteome Res. 2014;13(4):2109-19.
72. Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B, et al. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers n-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. Mol Cell Proteomics. 2014;13(7):1753-68.
73. Zhang F, Lin H, Gu A, Li J, Liu L, Yu T, et al. SWATH™- and iTRAQ-based quantitative proteomic analyses reveal an overexpression and biological relevance of cd109 in advanced nsclc. J Proteomics. 2014;102(0):125-36.
74. Sidoli S, Lin S, Xiong L, Bhanu NV, Karch KR, Johansen E, et al. SWATH analysis for characterization and quantification of histone post-translational modifications. Mol Cell Proteomics. 2015.
75. Arnhard K, Gottschall A, Pitterl F, Oberacher H. Applying 'sequential windowed acquisition of all theoretical fragment ion mass spectra' (SWATH) for systematic toxicological analysis with liquid chromatography-high-resolution tandem mass spectrometry. Anal Bioanal Chem. 2015;407(2):405-14.
76. Zhu X, Chen Y, Subramanian R. Comparison of information-dependent acquisition, SWATH, and msall techniques in metabolite identification study employing ultrahigh-performance liquid chromatography–quadrupole time-of-flight mass spectrometry. Anal Chem. 2014;86(2):1202-9.
77. Siegel D, Meinema AC, Permentier H, Hopfgartner G, Bischoff R. Integrated quantification and identification of aldehydes and ketones in biological samples. Anal Chem. 2014;86(10):5089-100.
78. Lemp MA, Beuerman RW. Cornea. In: Krachermer JA, Mannis MJ, Holland EJ, editors. Cornea. 3rd ed. St Louis: Mosby; 2011. p. 33-41.
79. Montés-Micó R. Role of the tear film in the optical quality of the human eye. J Cataract Refract Surg. 2007;33(9):1631-5.
80. Zhou L, Beuerman RW, Foo Y, Liu S, Ang LPK, Tan DTH. Characterisation of human tear proteins using high-resolution mass spectrometry. Ann Acad Med Singapore. 2006;35(6):400-7.
81. Nicholson JK, Lindon JC. Systems biology: Metabonomics. Nature. 2008;455(7216):1054-6.
82. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, et al. The human serum metabolome. PLoS ONE. 2011;6(2):e16957.
83. Wishart DS, Lewis MJ, Morrissey JA, Flegel MD, Jeroncic K, Xiong Y, et al. The human cerebrospinal fluid metabolome. J Chromatogr B. 2008;871(2):164-73.
84. Bertram HC, Eggers N, Eller N. Potential of human saliva for nuclear magnetic resonance-based metabolomics and for health-related biomarker identification. Anal Chem. 2009;81(21):9188-93.

85. Graça G, Duarte IF, Goodfellow BJ, Barros AS, Carreira IM, Couceiro AB, et al. Potential of NMR spectroscopy for the study of human amniotic fluid. Anal Chem. 2007;79(21):8367-75.
86. Benton HP, Wong DM, Trauger SA, Siuzdak G. Xcms2: Processing tandem mass spectrometry data for metabolite identification and structural characterization. Anal Chem. 2008;80(16):6382-9.
87. Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. J Chromatogr A. 2007;1158(1-2):318-28.
88. Want EJ, Coen M, Masson P, Keun HC, Pearce JTM, Reily MD, et al. Ultra performance liquid chromatography-mass spectrometry profiling of bile acid metabolites in biofluids: Application to experimental toxicology studies. Anal Chem. 2010;82(12):5282-9.
89. Wolters DA, Washburn MP, Yates JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem. 2001;73(23):5683-90.
90. Nakatsukasa M, Sotozono C, Shimbo K, Ono N, Miyano H, Okano A, et al. Amino acid profiles in human tear fluids analyzed by high-performance liquid chromatography and electrospray ionization tandem mass spectrometry. Am J Ophthalmol. 2011;In Press.
91. Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR, et al. Metlin: A metabolite mass spectral database. Ther Drug Monit. 2005;27(6):747-51.
92. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, et al. Massbank: A public repository for sharing mass spectral data for life sciences. J Mass Spectrom. 2010;45(7):703-14.
93. Ristimaa J, Gergov M, Pelander A, Halmesmäki E, Ojanperä I. Broad-spectrum drug screening of meconium by liquid chromatography with tandem mass spectrometry and time-of-flight mass spectrometry. Anal Bioanal Chem. 2010;398(2):925-35.
94. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, et al. The human urine metabolome. PLoS ONE. 2013;8(9):e73076.
95. Gika HG, Theodoridis GA, Wilson ID. Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabonomic analysis of zucker rat urine. J Sep Sci. 2008;31(9):1598-608.
96. Birkler RID, Støttrup NB, Hermannson S, Nielsen TT, Gregersen N, Bøtker HE, et al. A uplc-MS/MS application for profiling of intermediary energy metabolites in microdialysis samples--a method for high-throughput. J Pharm Biomed Anal. 2010;53(4):983-90.
97. Mora L, Hernández-Cázares AS, Aristoy M-C, Toldrá F. Hydrophilic interaction chromatographic determination of adenosine triphosphate and its metabolites. Food Chem. 2010;123(4):1282-8.
98. Kind T, Fiehn O. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. BMC Bioinformatics. 2006;7(1):234.
99. Kammerer B, Frickenschmidt A, Müller C, Laufer S, Gleiter C, Liebich H. Mass spectrometric identification of modified urinary nucleosides used as potential biomedical markers by LC–ITMS coupling. Anal Bioanal Chem. 2005;382(4):1017-26.
100. Okuda S, Yamada T, Hamajima M, Itoh M, Katayama T, Bork P, et al. Kegg atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res. 2008;36(suppl 2):W423-W6.

101. Xu Y, Heilier J-F, Madalinski G, Genin E, Ezan E, Tabet J-C, et al. Evaluation of accurate mass and relative isotopic abundance measurements in the ltq-orbitrap mass spectrometer for further metabolomics database building. Anal Chem. 2010;82(13):5490-501.
102. Moss SE, Klein R, Klein BEK. Prevalence of and risk factors for dry eye syndrome. Arch Ophthalmol. 2000;118(9):1264-8.
103. Chia E-M, Mitchell P, Rochtchina E, Lee AJ, Maroun R, Wang JJ. Prevalence and associations of dry eye syndrome in an older population: The Blue Mountains eye study. Clin Experiment Ophthalmol. 2003;31(3):229-32.
104. Viso E, Rodriguez-Ares MT, Gude F. Prevalence of and associated factors for dry eye in a Spanish adult population (the Salnes eye study). Ophthalmic Epidemiol. 2009;16(1):15-21.
105. Shimmura S, Shimazaki J, Tsubota K. Results of a population-based questionnaire on the symptoms and lifestyles associated with dry eye. Cornea. 1999;18(4):408-11.
106. Lee AJ, Lee J, Saw S-M, Gazzard G, Koh D, Widjaja D, et al. Prevalence and risk factors associated with dry eye symptoms: A population based study in indonesia. Br J Ophthalmol. 2002;86(12):1347-51.
107. Jie Y, Xu L, Wu YY, Jonas JB. Prevalence of dry eye among adult chinese in the beijing eye study. Eye. 2008;23(3):688-93.
108. Sullivan BD, Whitmer D, Nichols KK, Tomlinson A, Foulks GN, Geerling G, et al. An objective approach to dry eye disease severity. Invest Ophthalmol Vis Sci. 2010;51(12):6125-30.
109. Khanal S, Tomlinson A, McFadyen A, Diaper C, Ramaesh K. Dry eye diagnosis. Invest Ophthalmol Vis Sci. 2008;49(4):1407-14.
110. Versura P, Nanni P, Bavelloni A, Blalock WL, Piazzi M, Roda A, et al. Tear proteomics in evaporative dry eye disease. Eye. 2010;24(8):1396-402.
111. Nichols JJ, Green-Church KB. Mass spectrometry-based proteomic analyses in contact lens-related dry eye. Cornea. 2009;28(10):1109-17 10.097/ICO.0b013e3181a2ad81.
112. Ho SN. Intracellular water homeostasis and the mammalian cellular osmotic stress response. J Cell Physiol. 2006;206(1):9-15.
113. Michea L, Ferguson DR, Peters EM, Andrews PM, Kirby MR, Burg MB. Cell cycle delay and apoptosis are induced by high salt and urea in renal medullary cells. American Journal of Physiology - Renal Physiology. 2000;278(2):F209-F18.
114. Diebold Y, Calonge M, de Salamanca AE, Callejo S, Corrales RM, Sáez V, et al. Characterization of a spontaneously immortalized cell line (IOBA-NHC) from normal human conjunctiva. Invest Ophthalmol Vis Sci. 2003;44(10):4263-74.
115. Gilbard JP, Farris R, Santamaria J, Ii. Osmolarity of tear microvolumes in keratoconjunctivitis sicca. Arch Ophthalmol. 1978;96(4):677-81.
116. Unwin RD, Griffiths JR, Whetton AD. Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano LC-MS/MS. Nature Protocols. 2010;5(9):1574-82.
117. Chen L, Zhou L, Chan ECY, Neo J, Beuerman RW. Characterization of the human tear metabolome by LC–MS/MS. J Proteome Res. 2011;10(10):4876-82.
118. Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, et al. Metabolights—an open-access general-purpose repository for metabolomics studies and associated meta-data. Nucleic Acids Res. 2013;41(D1):D781-D6.
119. Corrales RM, Luo L, Chang EY, Pflugfelder SC. Effects of osmoprotectants on hyperosmolar stress in cultured human corneal epithelial cells. Cornea. 2008;27(5):574-9.

120. Garrett Q, Khandekar N, Shih S, Flanagan JL, Simmons P, Vehige J, et al. Betaine stabilizes cell volume and protects against apoptosis in human corneal epithelial cells under hyperosmotic stress. Exp Eye Res. 2013;108(0):33-41.
121. Hua X, Su Z, Deng R, Lin J, Li D-Q, Pflugfelder SC. Effects of l-carnitine, erythritol and betaine on pro-inflammatory markers in primary human corneal epithelial cells exposed to hyperosmotic stress. Curr Eye Res. 2015;40(7):657-67.
122. Chen W, Zhang X, Li J, Wang Y, Chen Q, Hou C, et al. Efficacy of osmoprotectants on prevention and treatment of murine dry eye. Invest Ophthalmol Vis Sci. 2013;54(9):6287-97.
123. Burg MB, Ferraris JD. Intracellular organic osmolytes: Function and regulation. J Biol Chem. 2008;283(12):7309-13.
124. Gallazzini M, Burg MB. What's new about osmotic regulation of glycerophosphocholine. Physiology. 2009;24(4):245-9.
125. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. Proteomexchange provides globally coordinated proteomics data submission and dissemination. Nat Biotech. 2014;32(3):223-6.
126. Vizcaíno JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, et al. The proteomics identifications (pride) database and associated tools: Status in 2013. Nucleic Acids Res. 2013;41(D1):D1063-D9.
127. Kregel KC. Invited review: Heat shock proteins: Modifying factors in physiological stress responses and acquired thermotolerance. J Appl Physiol. 2002;92(5):2177-86.
128. Li J, Ferraris JD, Yu D, Singh T, Izumi Y, Wang G, et al. Proteomic analysis of high nacl-induced changes in abundance of nuclear proteins. Physiological Genomics. 2012;44(21):1063-71.
129. Toivola DM, Strnad P, Habtezion A, Omary MB. Intermediate filaments take the heat as stress proteins. Trends Cell Biol.20(2):79-91.
130. Zachara NE, Hart GW. Cell signaling, the essential role of O-GlcNAc! Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 2006;1761(5-6):599-617.
131. Herzog R, Bender TO, Vychytil A, Bialas K, Aufricht C, Kratochwill K. Dynamic *o*-linked *n*-acetylglucosamine modification of proteins affects stress responses and survival of mesothelial cells exposed to peritoneal dialysis fluids. J Am Soc Nephrol. 2014;25(12):2778-88.
132. Cunha-Vaz JG. Diabetic retinopathy. Hackensack, NJ: World Scientific Publishing Company; 2011.
133. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract. 2010;87(1):4-14.
134. Wong TY, Cheung N, Tay WT, Wang JJ, Aung T, Saw SM, et al. Prevalence and risk factors for diabetic retinopathy: The Singapore malay eye study. Ophthalmology. 2008;115(11):1869-75.
135. Zheng Y, Lamoureux EL, Lavanya R, Wu R, Ikram MK, Wang JJ, et al. Prevalence and risk factors of diabetic retinopathy in migrant indians in an urbanized society in asia: The Singapore indian eye study. Ophthalmology. 2012;119(10):2119-24.
136. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care. 2012;35(3):556-64.
137. S. Tan G, Kamran Ikram M, Y. Wong T. Traditional and novel risk factors of diabetic retinopathy and research challenges. Curr Med Chem. 2013;20(26):3189-99.

138. Mohamed Q, Gillies MC, Wong TY. Management of diabetic retinopathy: A systematic review. JAMA: J A Med Assoc. 2007;298(8):902-16.
139. Beulens J, Patel A, Vingerling J, Cruickshank J, Hughes A, Stanton A, et al. Effects of blood pressure lowering and intensive glucose control on the incidence and progression of retinopathy in patients with type 2 diabetes mellitus: A randomised controlled trial. Diabetologia. 2009;52(10):2027-36.
140. ACCORD Study Group, ACCORD Eye Study Group. Effects of medical therapies on retinopathy progression in type 2 diabetes. N Engl J Med. 2010;363(3):233-44.
141. Pirola L, Balcerczyk A, Okabe J, El-Osta A. Epigenetic phenomena linked to diabetic complications. Nature Reviews Endocrinology. 2010;6(12):665-75.
142. Zavrelova H, Hoekstra T, Alssema M, Welschen LMC, Nijpels G, Moll AC, et al. Progression and regression: Distinct developmental patterns of diabetic retinopathy in patients with type 2 diabetes treated in the diabetes care system West-Friesland, the Netherlands. Diabetes Care. 2011;34(4):867-72.
143. Lamoureux EL, Wong TY. Diabetic retinopathy in 2011: Further insights from new epidemiological studies and clinical trials. Diabetes Care. 2011;34(4):1066-7.
144. Lehmann M, Yanes O, Krohne TU, Dorsey AL, Aguilar E, Marchetti V, et al. Metabolomic analysis of serum from diabetic patients with and without retinopathy. Association for Research in Vision and Opthalmology (ARVO) Meeting 2011; April 22, 2011; Fort Lauderdale, FL2011. p. A386.
145. Huang E-J, Kuo W-W, Chen Y-J, Chen T-H, Chang M-H, Lu M-C, et al. Homocysteine and other biochemical parameters in type 2 diabetes mellitus with different diabetic duration or diabetic retinopathy. Clin Chim Acta. 2006;366(1-2):293-8.
146. Yonem A, Duran C, Unal M, Ipcioglu OM, Ozcan O. Plasma apelin and asymmetric dimethylarginine levels in type 2 diabetic patients with diabetic retinopathy. Diabetes Res Clin Pract. 2009;84(3):219-23.
147. Xia J-F, Wang Z-H, Liang Q-L, Wang Y-M, Li P, Luo G-A. Correlations of six related pyrimidine metabolites and diabetic retinopathy in chinese type 2 diabetic patients. Clin Chim Acta. 2011;412(11-12):940-5.
148. ETDRS Research Group. Fundus photographic risk factors for progression of diabetic retinopathy. Etdrs report number 12. Early treatment diabetic retinopathy study research group. Ophthalmology. 1991;98(5):823-33.
149. Chan ECY, Pasikanti KK, Nicholson JK. Global urinary metabolic profiling procedures using gas chromatography-mass spectrometry. Nature Protocols. 2011;6(10):1483-99.
150. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.
151. Xia J, Broadhurst D, Wilson M, Wishart D. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. Metabolomics. 2013;9(2):280-99.
152. Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmüller E, et al. Gmd@csb.Db: The golm metabolome database. Bioinformatics. 2005;21(8):1635-8.
153. Xia J, Mandal R, Sinelnikov IV, Broadhurst D, Wishart DS. Metaboanalyst 2.0—a comprehensive server for metabolomic data analysis. Nucleic Acids Res. 2012.

154. Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, de Bono B, et al. Reactome knowledgebase of human biological pathways and processes. Nucleic Acids Res. 2009;37(suppl 1):D619-D22.
155. Pearl PL, Novotny EJ, Acosta MT, Jakobs C, Gibson KM. Succinic semialdehyde dehydrogenase deficiency in children and adults. Ann Neurol. 2003;54(S6):S73-S80.
156. Evans MD, Dizdaroglu M, Cooke MS. Oxidative DNA damage and disease: Induction, repair and significance. Mutation Research/Reviews in Mutation Research. 2004;567(1):1-61.
157. Brownlee M. The pathobiology of diabetic complications. Diabetes. 2005;54(6):1615-25.
158. Liu Y, Turdi S, Park T, Morris NJ, Deshaies Y, Xu A, et al. Adiponectin corrects high-fat diet–induced disturbances in muscle metabolomic profile and whole-body glucose homeostasis. Diabetes. 2013;62(3):743-52.
159. den Hartog GJM, Boots AW, Adam-Perrot A, Brouns F, Verkooijen IWCM, Weseler AR, et al. Erythritol is a sweet antioxidant. Nutrition. 2010;26(4):449-58.
160. Musso G, Gambino R, Cassader M. Obesity, diabetes, and gut microbiota: The hygiene hypothesis expanded? Diabetes Care. 2010;33(10):2277-84.
161. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, et al. Gut metagenome in european women with normal, impaired and diabetic glucose control. Nature. 2013;498(7452):99-103.
162. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature. 2012;490(7418):55-60.
163. Lees HJ, Swann JR, Wilson ID, Nicholson JK, Holmes E. Hippurate: The natural history of a mammalian–microbial cometabolite. J Proteome Res. 2013;12(4):1527-46.
164. Ferrannini E, Natali A, Camastra S, Nannipieri M, Mari A, Adam K-P, et al. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. Diabetes. 2013;62(5):1730-7.
165. Calder PC. Fatty acids and inflammation: The cutting edge between food and pharma. Eur J Pharmacol. 2011;668, Supplement 1(0):S50-S8.
166. Hirayama A, Nakashima E, Sugimoto M, Akiyama S-i, Sato W, Maruyama S, et al. Metabolic profiling reveals new serum biomarkers for differentiating diabetic nephropathy. Anal Bioanal Chem. 2012;404(10):3101-9.
167. Wallin A, Di Giuseppe D, Orsini N, Patel PS, Forouhi NG, Wolk A. Fish consumption, dietary long-chain n-3 fatty acids, and risk of type 2 diabetes: Systematic review and meta-analysis of prospective studies. Diabetes Care. 2012;35(4):918-29.
168. Mäkinen V-P, Tynkkynen T, Soininen P, Peltola T, Kangas AJ, Forsblom C, et al. Metabolic diversity of progressive kidney disease in 325 patients with type 1 diabetes (the FinnDiane study). J Proteome Res. 2011;11(3):1782-90.
169. Connor KM, SanGiovanni JP, Lofqvist C, Aderman CM, Chen J, Higuchi A, et al. Increased dietary intake of [omega]-3-polyunsaturated fatty acids reduces pathological retinal angiogenesis. Nat Med. 2007;13(7):868-73.
170. Sapieha P, Chen J, Stahl A, Seaward MR, Favazza TL, Juan AM, et al. Omega-3 polyunsaturated fatty acids preserve retinal function in type 2 diabetic mice. Nutr Diabetes. 2012;2:e36.
171. Tikhonenko M, Lydic TA, Opreanu M, Li Calzi S, Bozack S, McSorley KM, et al. N-3 polyunsaturated fatty acids prevent diabetic retinopathy by inhibition of retinal vascular damage and enhanced endothelial progenitor cell reparative function. PLoS ONE. 2013;8(1):e55177.

172. Fiehn O, Kopka J, Trethewey RN, Willmitzer L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. Anal Chem. 2000;72(15):3573-80.
173. Wagner C, Sefkow M, Kopka J. Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. Phytochemistry. 2003;62(6):887-900.
174. Chiang PPC, Lamoureux EL, Cheung CY, Sabanayagam C, Wong W, Tai ES, et al. Racial differences in the prevalence of diabetes but not diabetic retinopathy in a multi-ethnic asian population. Invest Ophthalmol Vis Sci. 2011;52(10):7586-92.
175. Rosenberger G, Koh CC, Guo T, Röst HL, Kouvonen P, Collins BC, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. Sci Data. 2014;1:e140031.
176. Yancey PH. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J Exp Biol. 2005;208(15):2819-30.
177. Zi J, Zhang S, Zhou R, Zhou B, Xu S, Hou G, et al. Expansion of the ion library for mining SWATH-MS data through fractionation proteomics. Anal Chem. 2014;86(15):7242-6.
178. Xu Y-F, Lu W, Rabinowitz JD. Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography–mass spectrometry-based metabolomics. Anal Chem. 2015;87(4):2273-81.
179. Denery JR, Nunes AAK, Dickerson TJ. Characterization of differences between blood sample matrices in untargeted metabolomics. Anal Chem. 2011;83(8):1040-7.
180. Mahieu NG, Huang X, Chen Y, Jr., Patti GJ. Credentialing features: A platform to benchmark and optimize untargeted metabolomic methods. Anal Chem. 2014;86(19):9583-9.

8. Appendices

8.1. Supplementary Methods in *In Vitro* Model of Dry Eye Disease

8.1.1. Cell Viability Assay

Three sets of cells were grown and treated in black 96-well plates. The average cell viability under each condition was determined with CellTiter-Glo Luminescent Cell Viability Assay (Promega, Madisin, WI, USA).

8.1.2. Sample Preparation for Global Metabonomic Analysis

Five sets of cells were grown in 6-well plates and treated with control, T1 and T2 treatment media. Cells were washed twice with cold PBS (4 °C) and lysed by the addition of 1.5 mL of methanol maintained at −80 °C. Plates were scraped over a dry ice bath, and each lysate was transferred to a microcentrifuge tube. Lysates were vortex-mixed at 1200 rpm for 5 min and centrifuged for 10 min at 16,000 g (4 °C). The supernatants were concentrated in a vacuum concentrator and freeze-dried subsequently.

8.1.3. Sample Preparation for Targeted Metabolite Analysis

Three sets of cells were grown in 6-well plates and treated with control, T1 and T2 treatment media. Cells were dissociated from the plates with TrypLE express (Invitrogen), collected in microcentifuge tubes and re-suspended in 500 µL of PBS. The cell suspension of each sample was divided into two aliquots of 450 uL and 50 uL respectively, and cells in each aliquot were pelleted.

Cells from the 50 uL aliquot were lysed in 40 uL of RIPA buffer containing 1× protease inhibitor (Thermo Scientific). Samples were vortexed briefly, placed in a cold ultrasonic bath for 30 s and mixed on an orbital shaker for 10 min. Lysates were centrifuged for 10 min at 15,000 g (4 °C) and the protein concentration of each supernatant was quantified with a DC protein assay kit (Bio-Rad Laboratories, Hercules, CA).

Cells from the 450 uL aliquot were lysed in 400 uL of 3:2 water/methanol containing 1 mM K_2EDTA and 10 mM ammonium bicarbonate in. Samples were vortexed briefly, snap-frozen in liquid nitrogen, thawed and homogenised with a hand-held homogeniser for 1 min. 600 uL of methanol was added and lysates were centrifuged for 10 min at 15,000 g (4 °C). The supernatants were dried and re-constituted in 30 uL of 9:1 methanol/water for targeted analysis. Metabolite levels in lysates were normalised to protein content.

8.1.4. Sample Preparation for Global Proteomic Analysis

Three sets of cells were grown in 6-well plates and treated with control, T1 and T2 treatment media. Cells were washed twice with cold PBS (4 °C) prior to harvesting. Flasks were scraped immediately after the addition of 1.5 mL of RIPA buffer with 1× protease inhibitor (Thermo Scientific). Cell suspensions were transferred to microcentrifuge tubes and gently agitated for 2 h at 4 °C to complete lysis. Lysates were centrifuged for 10 min at 16,000 g (4 °C) and the protein concentration of each supernatant was quantified using a Micro BCA protein assay kit (Thermo Scientific).

Cell lysates containing 50 µg of protein were reconstituted in 22 µL of 50 mM ammonium bicarbonate solution from the FASP kit denatured with 6 µL of 10% sodium dodecyl sulphate and 2 µL of tris(2-carboxyethyl) phosphine from the iTRAQ kit at 60^{o}C for 1 h. The reduced proteins were then transferred to 30 kD cut-off membrane cartridge from the FASP kit and centrifuged twice with 200 µL of urea solution from the FASP kit at 14,000 g for 15 min (25^{o}C). The reduced proteins were then incubated with 10 µL of 10× iodoacetamide from the FASP kit for 20 min in the dark and centrifuged twice at 14,000 g with 100 µL of urea solution, followed by ammonium bicarbonate solution. The samples were digested on-catridge with trypsin (AB Sciex) (1:25 trypsin/protein ratio) for 16 h at 37 °C. Digested peptides were eluted by centrifugation with 40 µL of ammonium bicarbonate solution, then with 50 µL of sodium chloride solution from the FASP kit. Digestion was quenched with 30 µL of 10% formic acid and samples were dried in a vacuum concentrator.

Samples were reconstituted in 20 µL of dissolution buffer from the iTRAQ kit. Each vial of iTRAQ reagent was reconstituted with 70 µL of ethanol. The respective iTRAQ reagents were then added to samples (one vial per 50 µg protein) and mixed for 2 hours at 25 °C to label the peptides. The labelled samples were then pooled together, dried and desalted on ultramicro spin columns (The Nest Group) prior to nano-LC-MS/MS analysis.

8.1.5. Peak Finding and Alignment of Metabonomic Data

Table 8.1. MarkerView parameters for peak finding and alignment of LC-MS/MS data obtained from IOBA-NHC lysates.

Data Processing Step	Parameters	Parameters	
		Positive Ionisation	Negative Ionisation
Peak detection	Subtraction offset/scans	3	
	Multiplication factor	1.2	
	Minimum spectral width/Da	0.01	
	Minimum chromatographic width/scans	3	
	RT window/min	0.6 to 18	
	Noise threshold/counts	20	5
Peak alignment	RT tolerance/min	0.15	0.20
	Mass tolerance/ppm	10	
	Minimum Number of Samples	5	
	Minimum Threshold	20	5

8.1.6. Targeted Metabonomic Analysis

Table 8.2. Optimised MRM parameters for validation of metabolite markers.

Metabolite	Q1 (m/z)	Q3 (m/z)	DP (V)	CE (V)	CXP (V)
Carnitine	162.1	103.0	60	23	12
Glycerophosphocholine	258.1	104.0	70	22	9
Phosphocholine	184.0	86.0	60	23	9

8.2. Supplementary Results for Metabotyping of Diabetic Retinopathy

Table 8.3 Comparison of clinicopathological characteristics in subset of participants with HbA1c levels at 10% or below.

Clinical Characteristics	DR	Non-DR	*P*-value
N	30	39	-
Gender	12 M	22 M	0.176†
Age (years)	61 (54–69)	62 (52–69)	0.807‡
Diastolic BP (mm Hg)	72 (67–79)	77 (68–83)	0.153‡
Systolic BP (mm Hg)	140 (131–156)	133 (122–146)	0.206‡
BMI (kg/m^2)	25.3 (23.1–29.4)	28.7 (23.3–30.3)	0.186‡
Serum creatinine (μM)	84 (69–125)	79 (58–124)	0.181‡
Blood glucose, random (mM)	9.3 (6.5–11.9)	8.2 (6.9–11.0)	0.599‡
Total cholesterol (mM)	4.42 (3.93–5.32)	4.32 (3.64–5.19)	0.498‡
LDL cholesterol (mM)	2.73 (2.15–3.44)	2.56 (2.07–3.27)	0.603‡
HDL cholesterol (mM)	1.01 (0.84–1.25)	1.01 (0.83–1.23)	0.950‡
Triglycerides (mM)	1.79 (1.30–2.54)	1.32 (0.93–2.07)	0.121‡
Blood HbA1c (%)	7.9 (7.1–8.8)	7.3 (6.5–8.0)	0.055‡
Duration of diabetes (years)	13 (7–22)	11 (6–20)	0.364‡
Urine ACR (μg/mg creatinine)	46.0 (14.6–123.8)	24.8 (12.3–57.0)	0.156‡
eGFR (mL/min/1.73 m^2)	71.0 (43.8–95.2)	87.1 (48.4–99.9)	0.196‡
Microalbuminuria	14	14	0.336†
Macroalbuminuria	4	2	0.230†
Chronic Kidney Disease§	12	11	0.303†
On Diabetes Medication	25	32	0.889†

* Values indicate median with interquartile range, where applicable
† Pearson chi-square test
‡ Mann-Whitney U test
§ Defined as eGFR below 60 mL/1.73 m^2

Table 8.4 Comparison of clinicopathological characteristics in subset of participants with HbA1c levels between 6 to 10%.

Clinical Characteristics	DR	Non-DR	*P*-value
N	29	34	-
Gender	11 M	19 M	0.155†
Age (years)	61 (55–69)	61 (52–67)	0.385‡
Diastolic BP (mm Hg)	72 (67–78)	75 (68–81)	0.124‡
Systolic BP (mm Hg)	141 (131–156)	133 (122–145)	0.097‡
BMI (kg/m^2)	25.6 (23.1–29.5)	28.6 (23.3–30.2)	0.224‡
Serum creatinine (µM)	85 (69–129)	77 (58–104)	0.126‡
Blood glucose, random (mM)	9.6 (6.5–12.1)	8.5 (7.0–11.2)	0.719‡
Total cholesterol (mM)	4.43 (4.01–5.33)	4.22 (3.69–5.32)	0.500‡
LDL cholesterol (mM)	2.78 (2.16–3.45)	2.54 (2.07–3.12)	0.508‡
HDL cholesterol (mM)	1.03 (0.83–1.26)	1.00 (0.86–1.24)	0.915‡
Triglycerides (mM)	1.79 (1.38–2.57)	1.35 (0.94–2.33)	0.151‡
Blood HbA1c (%)	7.9 (7.1–8.8)	7.4 (6.8–8.1)	0.161‡
Duration of diabetes (years)	13 (7–22)	10 (6–18)	0.236‡
Urine ACR (µg/mg creatinine)	46.0 (17.7–129.1)	26.1 (11.5–52.8)	0.086‡
eGFR (mL/min/1.73 m^2)	63.3 (42.4–94.7)	88.1 (55.2–100.3)	0.095‡
Microalbuminuria	14	12	0.216†
Macroalbuminuria	4	2	0.286†
Chronic Kidney Disease§	12	9	0.211†
On Diabetes Medication	24	30	0.536†

* Values indicate median with interquartile range, where applicable
† Pearson chi-square test
‡ Mann-Whitney U test
§ Defined as eGFR below 60 mL/1.73 m^2

Table 8.5 Odd ratios metabolite markers adjusted for age, gender and diastolic BP.

Metabolite	Odds ratio, per s.d. increment in metabolite peak area (95% CI)			
	Basic Model	+Age	+ Gender	+ Diastolic BP
1,5-Anhydroglucitol	0.37 (0.19–0.65)*	0.37 (0.19–0.65)*	0.36 (0.18–0.63)*	0.30 (0.14–0.54)*‡
1,5-Gluconolactone	2.40 (1.38–4.65)*	2.50 (1.42–4.95)*	2.49 (1.40–4.94)*	2.50 (1.43–4.88)*
2-Deoxyribonic acid	2.08 (1.25–3.76)*	2.17 (1.29–3.98)*	2.08 (1.25–3.76)*	2.15 (1.29–3.91)*
3,4-Dihydroxybutyric acid	1.61 (1.01–2.70)†	1.71 (1.05–2.92)*	1.61 (1.01–2.69)†	1.63 (1.02–2.73)*
3-Methyl-2-oxovaleric acid	0.81 (0.50–1.27)	0.72 (0.43–1.19)	0.81 (0.49–1.29)	0.78 (0.48–1.24)
α-Ketoisovalerate	0.58 (0.34–0.93)*	0.53 (0.30–0.88)*	0.58 (0.34–0.93)*	0.55 (0.32–0.90)*
Erythritol	1.68 (1.04–2.85)*	1.74 (1.07–2.98)*	1.68 (1.04–2.86)*	1.73 (1.08–2.94)*
Erythronic acid	1.57 (0.97–2.71)†	1.65 (1.00–2.90)†	1.57 (0.97–2.71)†	1.59 (0.98–2.75)†
Gluconic acid	2.65 (1.50–5.33)*	2.64 (1.49–5.32)*	2.66 (1.50–5.37)*	2.79 (1.57–5.61)*
Glutamine	0.51 (0.28–0.86)*	0.51 (0.28–0.85)*	0.51 (0.28–0.86)*	0.53 (0.30–0.88)*
Hippuric acid	1.71 (1.05–2.98)*	1.94 (1.14–3.61)*	1.72 (1.05–3.00)*	1.74 (1.08–3.02)*
Inositol phosphate	1.26 (0.80–2.03)	1.28 (0.81–2.08)	1.26 (0.79–2.07)	1.32 (0.84–2.14)
Lactose/Cellobiose	1.78 (1.10–3.01)*	1.88 (1.15–3.25)*	1.79 (1.10–3.04)*	1.85 (1.14–3.14)*
Maltose/Trehalose	2.14 (1.29–3.84)*	2.13 (1.28–3.82)*	2.18 (1.31–3.94)*	2.33 (1.39–4.20)*
Mannose	1.81 (1.09–3.27)*	1.82 (1.07–3.38)*	1.86 (1.10–3.44)*	1.90 (1.14–3.47)*
Myo-inositol	1.60 (0.99–2.76)†	1.70 (1.03–3.02)*	1.60 (0.99–2.76)†	1.69 (1.04–2.96)*
Ribose	1.65 (1.02–2.80)*	1.71 (1.05–2.95)*	1.67 (1.03–2.85)*	1.74 (1.09–2.95)*
Urea	1.95 (1.18–3.48)*	2.28 (1.32–4.30)*	1.95 (1.18–3.48)*	2.05 (1.22–3.71)*
* P(model)< 0.05 † P(model)<0.1	‡ P(likelihood-ratio)<0.05 § P(likelihood-ratio)<0.1			

Table 8.6 Odd ratios metabolite markers adjusted for total cholesterol, LDL cholesterol and HDL cholesterol.

Metabolite	Odds ratio, per s.d. increment in metabolite peak area (95% CI)			
	Basic Model	+ Total Cholesterol	+LDL Cholesterol	+ HDL Cholesterol
1,5-Anhydroglucitol	0.37 (0.19–0.65)*	0.36 (0.18–0.63)*	0.36 (0.19–0.63)*	0.35 (0.18–0.61)*
1,5-Gluconolactone	2.40 (1.38–4.65)*	2.58 (1.47–5.01)*	2.62 (1.50–5.12)*	2.55 (1.47–4.93)*
2-Deoxyribonic acid	2.08 (1.25–3.76)*	2.21 (1.33–4.00)*	2.30 (1.37–4.20)*	2.20 (1.32–3.99)*
3,4-Dihydroxybutyric acid	1.61 (1.01–2.70)†	1.65 (1.04–2.76)*	1.68 (1.05–2.80)*	1.67 (1.05–2.78)*
3-Methyl-2-oxovaleric acid	0.81 (0.50–1.27)	0.73 (0.45–1.15)	0.73 (0.45–1.15)	0.75 (0.47–1.18)
α-Ketoisovalerate	0.58 (0.34–0.93)*	0.51 (0.29–0.84)*	0.52 (0.30–0.84)*	0.54 (0.31–0.87)*
Erythritol	1.68 (1.04–2.85)*	1.80 (1.12–3.06)*	1.85 (1.14–3.15)*	1.78 (1.11–3.00)*
Erythronic acid	1.57 (0.97–2.71)†	1.70 (1.04–2.96)*	1.74 (1.07–3.06)*	1.65 (1.02–2.83)†
Gluconic acid	2.65 (1.50–5.33)*	2.73 (1.54–5.47)*	2.75 (1.55–5.53)*	2.79 (1.57–5.61)*
Glutamine	0.51 (0.28–0.86)*	0.55 (0.31–0.92)*	0.55 (0.30–0.91)*	0.51 (0.28–0.86)*
Hippuric acid	1.71 (1.05–2.98)*	1.82 (1.11–3.19)*	1.85 (1.13–3.28)*	1.77 (1.10-3.08)*
Inositol phosphate	1.26 (0.80–2.03)	1.44 (0.91–2.35)	1.43 (0.90–2.32)	1.36 (0.87–2.18)
Lactose/Cellobiose	1.78 (1.10–3.01)*	1.92 (1.19–3.28)*	1.98 (1.22–3.39)*	1.90 (1.18–3.20)*
Maltose/Trehalose	2.14 (1.29–3.84)*	2.04 (1.25–3.59)*	2.05 (1.25–3.61)*	2.10 (1.28–3.68)*
Mannose	1.81 (1.09–3.27)*	1.64 (1.00–2.93)†	1.66 (1.00–2.96)†	1.71 (1.05–3.04)*
Myo-inositol	1.60 (0.99–2.76)†	1.75 (1.08–3.03)*	1.80 (1.11–3.13)*	1.74 (1.08–3.00)*
Ribose	1.65 (1.02–2.80)*	1.69 (1.05–2.89)*	1.71 (1.07–2.93)*	1.71 (1.07–2.90)*
Urea	1.95 (1.18–3.48)*	2.19 (1.32–3.92)*	2.23 (1.34–4.00)*	2.05 (1.25–3.64)*
* P(model)< 0.05 † P(model)<0.1	‡ P(likelihood-ratio)<0.05 § P(likelihood-ratio)<0.1			